

STIC-Biotech/ChemLib

138619

From: Chan, Christina
Sent: Tuesday, November 16, 2004 4:33 PM
To: Walicka, Malgorzata
Cc: STIC-Biotech/ChemLib
Subject: RE:

Please rush. Thanks Chris

Chris Chan
TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841
Remsen, 3E89

-----Original Message-----

From: Walicka, Malgorzata
Sent: Tuesday, November 16, 2004 4:19 PM
To: Chan, Christina
Cc: STIC-Biotech/ChemLib
Subject:

Christina,

Please authorize my RUSH search of SEQ ID NO:8 in 10/802,682. Thank you.

Malgorzata A. Walicka, Ph.D.
Patent Examiner
Art Unit 1652, Recombinant Enzymes
USPTO, Remsen Building, Room 2C76
400 Dulany St.
Alexandria, VA 22313
Mail Room 2C70
Tel. (571)272-0944, fax (571)273-0944

RECEIVED
NOV 16 2004
STIC-BIOTECH/CHEM LIB
(STIC)

STAFF USE ONLY

Searcher: P. Schneider
Searcher Phone: 2- 2526
Date Searcher Picked up: 11/18
Date Completed: 11/18
Searcher Prep/Rev. Time: 11
Online Time: 5

Type of Search

NA Sequence: # _____
AA Sequence: # 1 _____
Structure: # _____
Bibliographic: _____
Litigation: _____
Patent Family: _____
Other: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: Compuser
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)

acid by standard procedures. The derivatives of ADH enzymes have desired substrate specificity, higher affinity to a substrate, lower affinity to an inhibitory compound, higher stability against temperature and/or pH and higher catalytic speed

Sequence 579 AA:

Query Match 100.0%; Score 3069; DB 2; Length 579;
Best Local Similarity 100.0%; Pred. No. 4,3e-244;
Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MNPPTLTSAVLLTPAPAFQVPTITDELLANPAGEMINYGROENYRHSPLTQIT 60
1 MNPPTLTSAVLLTPAPAFQVPTITDELLANPAGEMINYGROENYRHSPLTQIT 60
61 ADVNGQLVWAKMEAGAVQVTPMIDGVTLANPBDVIOALDAQTGLIWEHRQLPA 120
61 ADVNGQLVWAKMEAGAVQVTPMIDGVTLANPBDVIOALDAQTGLIWEHRQLPA 120
121 VATLNAQGRKRGVALVGTSLYFSSWMDNLIALDMETGVVFPVERSGSGDGLTSNTGP 180
121 VATLNAQGRKRGVALVGTSLYFSSWMDNLIALDMETGVVFPVERSGSGDGLTSNTGP 180
121 VATLNAQGRKRGVALVGTSLYFSSWMDNLIALDMETGVVFPVERSGSGDGLTSNTGP 180
181 IVANGVIVAGSTCOYSPYGCPSIGHDSATGEBELMRNHFIPQGEDEGTWGNDEARMT 240
181 IVANGVIVAGSTCOYSPYGCPSIGHDSATGEBELMRNHFIPQGEDEGTWGNDEARMT 240
241 GWMQGITPTNTNLYFGSTGVGPASRGTGRTGGLTYGNTNFAVRPDGELIWEHRQL 300
241 GWMQGITPTNTNLYFGSTGVGPASRGTGRTGGLTYGNTNFAVRPDGELIWEHRQL 300
301 PRDNDQCEFTFEMVAVNDVQPSAEMGLRANINNAATGERRVLTGAPCKTGWMSFDA 360
301 PRDNDQCEFTFEMVAVNDVQPSAEMGLRANINNAATGERRVLTGAPCKTGWMSFDA 360
361 SGEFLMARDNTNMTASIDETGLVTNEDAVLKELDVEYDVCPTFGGDMSSAALNP 420
361 SGEFLMARDNTNMTASIDETGLVTNEDAVLKELDVEYDVCPTFGGDMSSAALNP 420
421 TGIYFLPLNNAACVIMAVDOEFSAIDVYNTSATKLAPEGEMGRIDAIDISTRTLMSA 480
421 TGIYFLPLNNAACVIMAVDOEFSAIDVYNTSATKLAPEGEMGRIDAIDISTRTLMSA 480
481 ERPAANYSPLSTAGVNGGTDYFRALSOETGETLQARLATVATGQALSYELDGVQ 540
481 ERPAANYSPLSTAGVNGGTDYFRALSOETGETLQARLATVATGQALSYELDGVQ 540
541 YIAGAGLTYGTOLNAPLAFAIDSTSVGNAIVYFALPQ 579
541 YIAGAGLTYGTOLNAPLAFAIDSTSVGNAIVYFALPQ 579

RESULT 2
ADN10956
ID ADN10956 standard; protein; 579 AA.

XX AC ADN10956;

XX ADN10956;

XX 01-JUL-2004 (first entry)

XX Glucobacter oxydans Enzyme B, used in ascorbic acid production.

XX Enzyme B; ascorbic acid; vitamin C; L-gulonol-1,4-lactone; L-gulonic acid;

XX L-galactono-1,4-lactone; L-galactonic acid.

XX Glucobacter oxydans.

XX WO2004029267-A1.

XX 08-APR-2004.

XX 22-SEP-2003; 2003WO-EP010489.

PR 27-SEP-2002; 2002EP-00021602.

XX (STM) DSM IP ASSETS BV.

XX Hoshino T, Shinjoh M;

XX WPI; 2004-329889/30.

XX N-PSDB; ADN10955.

PT Producing L-ascorbic acid using enzyme B of Glucobacter oxydans, from

PT Substrates L-gulose, L-galactose, L-idose, and L-talose.

PS Claim 1; SEQ ID NO 2; 24pp; English.

The present sequence is the protein sequence of Enzyme B from Glucobacter oxydans strain DSM 4025. Enzyme B has a molecular weight of about 60,000 Da by SDS-PAGE, substrate specificity for primary and secondary alcohols and aldehydes, is stable in the pH range 6-9 with optimal activity at about pH 8.0, and is inhibited by Cu²⁺, Zn²⁺, Mn²⁺, Fe²⁺ and Fe³⁺. The present invention provides the use of this enzyme in a process for producing L-ascorbic acid from L-gulose, L-galactose, L-idose or L-talose, or from L-gulonol-1,4-lactone, L-gulonic acid, L-galactono-1,4-lactone, L-galactonic acid, L-idono-1,4-lactone, L-idonic acid, L-talono-1,4-lactone and L-talonic acid. Enzyme B is also used in a process for the production of L-gulonol-1,4-lactone or L-gulonic acid from L-gulose, and L-galactono-1,4-lactone or L-galactonic acid from L-galactose. The processes involve contacting the enzyme with the respective substrate and isolating the product from the reaction mixture. The process is conducted for 1-20 hours at pH 1-9 (preferably pH 2-8) and 13-45 (preferably 18-42) degrees C. Production of L-gulonol-1,4-lactone/L-gulonic acid from L-gulose, vitamin C from L-gulonol-1,4-lactone/L-gulonic acid, L-galactono-1,4-lactone/L-galactonic acid from L-galactose, and vitamin C from L-galactono-1,4-lactone/L-galactonic acid by recombinant *Escherichia coli* JM109 carrying the Enzyme B gene is described in examples from the invention.

XX Sequence 579 AA:

Query Match 100.0%; Score 3069; DB 8; Length 579;
Best Local Similarity 100.0%; Pred. No. 4,3e-244;
Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MNPPTLTSAVLLTPAPAFQVPTITDELLANPAGEMINYGROENYRHSPLTQIT 60
1 MNPPTLTSAVLLTPAPAFQVPTITDELLANPAGEMINYGROENYRHSPLTQIT 60
61 ADVNGQLVWAKMEAGAVQVTPMIDGVTLANPBDVIOALDAQTGLIWEHRQLPA 120
61 ADVNGQLVWAKMEAGAVQVTPMIDGVTLANPBDVIOALDAQTGLIWEHRQLPA 120
121 VATLNAQGRKRGVALVGTSLYFSSWMDNLIALDMETGVVFPVERSGSGDGLTSNTGP 180
121 VATLNAQGRKRGVALVGTSLYFSSWMDNLIALDMETGVVFPVERSGSGDGLTSNTGP 180
121 VATLNAQGRKRGVALVGTSLYFSSWMDNLIALDMETGVVFPVERSGSGDGLTSNTGP 180
181 IVANGVIVAGSTCOYSPYGCPSIGHDSATGEBELMRNHFIPQGEDEGTWGNDEARMT 240
181 IVANGVIVAGSTCOYSPYGCPSIGHDSATGEBELMRNHFIPQGEDEGTWGNDEARMT 240
241 GWMQGITPTNTNLYFGSTGVGPASRGTGRTGGLTYGNTNFAVRPDGELIWEHRQL 300
241 GWMQGITPTNTNLYFGSTGVGPASRGTGRTGGLTYGNTNFAVRPDGELIWEHRQL 300
301 PRDNDQCEFTFEMVAVNDVQPSAEMGLRANINNAATGERRVLTGAPCKTGWMSFDA 360
301 PRDNDQCEFTFEMVAVNDVQPSAEMGLRANINNAATGERRVLTGAPCKTGWMSFDA 360
361 SGEFLMARDNTNMTASIDETGLVTNEDAVLKELDVEYDVCPTFGGDMSSAALNP 420
361 SGEFLMARDNTNMTASIDETGLVTNEDAVLKELDVEYDVCPTFGGDMSSAALNP 420
421 TGIYFLPLNNAACVIMAVDOEFSAIDVYNTSATKLAPEGEMGRIDAIDISTRTLMSA 480
421 TGIYFLPLNNAACVIMAVDOEFSAIDVYNTSATKLAPEGEMGRIDAIDISTRTLMSA 480

QY 481 ERPAANTSPVISTAGVVFNGGTDRTYFRALSOETGETLMQARLATVATGQALSYELDVQ 540
 DB 481 ERPAANTSPVISTAGVVFNGGTDRTYFRALSOETGETLMQARLATVATGQALSYELDVQ 540
 QY 541 YTAIGAGGLTYGTQLNAPLAELAIDSTSVGNAYVFALPQ 579
 DB 541 YTAIGAGGLTYGTQLNAPLAELAIDSTSVGNAYVFALPQ 579

RESULT 3
 AD134121
 ID AD134121 standard; protein; 580 AA.
 AC AD134121;
 XX
 DE 15-APR-2004 (first entry)
 XX
 DE Ketoglulonicigenium sp. sorbitol dehydrogenase (SDH)3 protein.
 XX
 KM Sorbitol dehydrogenase; SDH; cytochrome c; molecular biology;
 KM bacteriology; industrial fermentation; L-sorbose; 2-keto-L-gulonic acid;
 KM 2KUG; enzyme.
 XX
 OS Ketoglulonicigenium.
 XX
 XX Key Location/Qualifiers
 FT Peptide 1..23
 FT /label= Signal_peptide
 FT Protein 24..580
 FT /note= "Mature SDH protein"

US2003228672-A1.
 11-DEC-2003.
 06-JUN-2002; 2002US-00162713.
 06-JUN-2002; 2002US-00162713.
 06-JUN-2002; 2002US-00162713.
 (ARCH) ARCHER-DANIELS MIDLAND CO.
 Choi E, D'elja J, Kim H, Kim M, Lee JK, Pan J, Stoddard SF,
 Yum D;
 MPI; 2004-052025/05.
 DR N-PSDB; AD134117, AD134124.
 XX
 FT New proteins of sorbitol dehydrogenases and cytochrome c of the strains
 FT Ketoglulonicigenium spp., useful in molecular biology, bacteriology and
 FT industrial fermentation, e.g. for producing L-sorbose and 2-keto-L-
 FT gulonic acid.
 XX
 PS Claim 79; SEQ ID NO 8; 68pp; English.
 XX
 CC The invention relates to the identification and isolation of nucleic acid
 CC sequences and proteins of sorbitol dehydrogenases (SDH) and cytochrome c
 CC of the strains, Ketoglulonicigenium sp. The proteins and nucleic acid
 CC molecules are useful in the fields of molecular biology, bacteriology and
 CC industrial fermentation specifically for producing L-sorbose and 2-keto-L-
 CC -gulonic acid (2KUG). The present sequence is Ketoglulonicigenium sp. SDH
 CC protein.
 CC
 SQ Sequence 580 AA;
 Query Match 90.1%; Score 2765; DB 8; Length 580;
 Best Local Similarity 87.6%; Pred. No. 5.2e-219;
 Matches 507; Conservative 38; Mismatches 34; Indels 0; Gaps 0;

QY 1 MNPTTLTSAVALLTTPAFAQVTPITDELNPAGGINTGRQENRHSPLQIT 60
 DB 1 MRPTTLTSAVALLTTPAFAQVTPITDELNPAGGINTGRQENRHSPLQIT 60

QY 61 ADNVGLOLVARMEAGAVQVTPMINDGVXYLIANPGVIOALDAQTGLIMEHRRLA 120
 DB 61 TTNVQGLQVMARMEAGAVQVTPMINDGVXYLIANPGVIOALDAQTGLIMEHRRLA 120
 QY 121 VATLNAGDRKRGVALLYGTSLYSSMDNHLALDMETQGVVDVERSGEDDLTNTTGP 180
 DB 121 VASINGQDRKRGVALLYGTSLYSSMDNHLALDMETQGVVDVERSGEDDLTNTTGP 180
 QY 181 IVANGVYVAGSTCGQSPVCGPCTISGHD SATGEELMKNHFIPOGEGEDTWNDFEARMYT 240
 DB 181 IVANGVYVAGSTCGQSPVCGPCTISGHD SATGEELMKNHFIPOGEGEDTWNDFEARMYT 240
 QY 241 GVMGQITVDPTNVLVYFGSTGVGPASEORGTGPGTLVGTNRFAVRPTGTGIYVRHQT 300
 DB 241 GVMGQITVDPTNVLVYFGSTGVGPASEORGTGPGTLVGTNRFAVRPTGTGIYVRHQT 300
 QY 301 PRDNMDQECTFEMVAVNDVQPSAMEGLRALNPAAATGERRVLTGAPCKTGTWSPDA 360
 DB 301 PRDNMDQECTFEMVAVNDVQPSAMEGLRALNPAAATGERRVLTGAPCKTGTWSPDA 360
 QY 361 SGEFLMARPTNTNMIASTIDETGLTVNEDAVLXELDYEDVCPFLGSRDMSAALPD 420
 DB 361 TGEFLMARPTNTNMIASTIDETGLTVNEDAVLXELDYEDVCPFLGSRDMSAALPD 420
 QY 421 TGIYFPLNNACVDIMAVDQEFSAIDVYNTSATAKIAGFENMGRIDAIDISTGRTLWSA 480
 DB 421 SGIFYFPLNNACVDIMAVDQEFSAIDVYNTSATAKIAGFENMGRIDAIDISTGRTLWSA 480
 QY 481 ERPAANTSPVISTAGVVFNGGTDRTYFRALSOETGETLMQARLATVATGQALSYELDVQ 540
 DB 481 ERPAANTSPVISTAGVVFNGGTDRTYFRALSOETGETLMQARLATVATGQALSYELDVQ 540
 QY 541 YTAIGAGGLTYGTQLNAPLAELAIDSTSVGNAYVFALPQ 579
 DB 541 YTAIGAGGLTYGTQLNAPLAELAIDSTSVGNAYVFALPQ 579

RESULT 4
 AAM37873
 ID AAM37873 standard; protein; 579 AA.
 AC AAM37873;
 XX
 DE 10-AUG-1998 (first entry)
 XX
 DE Alcohol and/or aldehyde dehydrogenase A amino acid sequence.
 XX
 KM Alcohol/aldehyde dehydrogenase A enzyme; recombinant organism; aldehyde;
 KM ketone; carboxylic acid; L-sorbose; D-sorbitol; 2-keto-L-gulonic acid;
 KM L-ascorbic; inhibition.
 XX
 OS Gluconobacter oxydans.
 XX
 XX Key Location/Qualifiers
 FT Peptide 1..23
 FT /note= "signal peptide"
 FT Protein 24..579
 FT /note= "mature protein"

EP832974-A2.
 01-APR-1998.
 11-SEP-1997; 97EP-00115801.
 19-SEP-1996; 96EP-00115001.
 (HOFF) HOFFMANN LA ROCHE & CO AG F.
 Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
 WPI; 1998-195228/18.
 N-PSDB; AAV29051.

XX Recombinant Gluconobacter oxydans alcohol and/or aldehyde dehydrogenase
PT enzyme(s) - useful for converting substrate(s), e.g. L-sorbose or D-
PT sorbitol to 2-keto-L-gulononic acid.
XX
XX
XX

PS Claim 1; Page 35-37; 59pp; English.

XX This is the amino acid sequence for the Gluconobacter oxydans alcohol
CC and/or aldehyde dehydrogenase A enzyme. The enzymes or recombinant
CC organisms can be used to convert suitable substrates to aldehydes,
CC ketones or carboxylic acids, especially to convert L-sorbose or D-
CC sorbitol to 2-keto-L-gulononic acid, which can be converted to L-ascorbic
CC acid by standard procedures. The derivatives of ADH enzymes have desired
CC substrate specificity, higher affinity to a substrate, lower affinity to
CC an inhibitory compound, higher stability against temperature and/or pH
CC and higher catalytic speed
XX
XX

SO Sequence 579 AA;

Query Match 85.1%; Score 2611; DB 2; Length 579;
Best Local Similarity 82.1%; Pred. No. 2,7e-206;
Matches 476; Conservative 51; Mismatches 51; Indels 2; Gaps 2;

QY 1 MNPTTLRTSAVLLTAPAAFAQVPTITDELLANPPAGEMWINGRQENYRHSPLTQIT 60
DB 1 MKPTSLMASAGALALAAAPAFQVPTITDELLANPPAGEMWISYGRQENYRHSPLTQIT 60
QY 61 ADVNGQLQVWARGMEAGAVQVTPMIDGVWYLANPQDVTOALDACTGDLIWEHRQLPA 120
DB 61 TENNGQLQVWARGMEAGAVQVTPMIDGVWYLANPQDVTOALDACTGDLIWEHRQLPA 120
QY 121 VATNAGDGRKRGVALGTSLYFSWMNHLIALDMRGVAVPVERSGEGDGLTSNTGP 180
DB 121 IATNLSGEBPRGALGTVNTPVSWDNLVALDPAQTGQVTPVDRGQSD -NVSNSGP 179
QY 181 IVANGVIVAGSTCOYSPYGCFTSGHDSATGEELRNHFIPOGEBGDETWNDFEARWMT 240
DB 180 IVANGVIVAGSTCOYSPYGCFTSGHDSATGEELRNHFIPOGEBGDETWNDFEARWMT 239
QY 241 GWMQGITVDPVTNLVFGYSTGVGASSETGRTGCGTLYGNTTFEAVRPTDGEIVMHHQTL 300
DB 240 GWMQGITVDPVTNLVFGYSTGVGASSETGRTGCGTLYGNTTFEAVRPTDGEIVMHHQTL 299
QY 301 PRDWMDOECTEFEMVAVNDVOPSAEMEGELAINPNAATGERRVLTGAPCKTGTWMSFDA 360
DB 300 PRDWMDOECTEFEMVAVNDVOPSAEMEGELAINPNAATGERRVLTGAPCKTGTWMSFDA 359
QY 361 SGEFLNARDNYTMTASIDETGLVTNEDAVLKELDVEYDVCPFLGGRDMSAALNDP 420
DB 360 TGEFLNARDNYTMTASIDETGLVTNEDAVLKELDVEYDVCPFLGGRDMSAALNDP 419
QY 421 TGIYFLPLNNACYIMAVDOEFASLDVYNTSATKALPFEENGRIDAIDISTGRTLMSA 480
DB 420 SGIYFLPLNNACYIMAVDOEFASLDVYNTSATKALPFEENGRIDAIDISTGRTLMSA 479
QY 481 ERPAANTSPVLTAGSVFENGTDTRYFRALSQETGETLMOARLATYATGCAIYEILDQV 540
DB 480 ERAAANTSPVLTAGSVFENGTDTRYFRALSQETGETLMOARLATYATGCAIYEILDQV 539
QY 541 YIAIAGAGLTGTQOLNAPLA-BAIDSTSVGNATVVFALPQ 579
DB 540 YVAILAGGVSIGSGLNSALAGERVDTAIGNAVVVFALPQ 579

RESULT 5
AD134118
ID AD134118 standard; protein; 578 AA.

XX AD134118;
XX 15-APR-2004 (first entry)
XX Ketogulononicigenium sp. sorbitol dehydrogenase (SDH) 1 protein.
DE

XX Sorbitol dehydrogenase; SDH; cytochrome c; molecular biology;
KW bacteriology; industrial fermentation; L-sorbose; 2-keto-L-gulononic acid;
KW 2KLG; enzyme.
XX
XX
XX

OS Ketogulononicigenium.

FT Key Location/Qualifiers
FT Peptide 1..23
FT Protein 24..578
FT /note= "Mature SDH protein"

UN2003228672-A1.

11-DEC-2003.

06-JUN-2002; 2002US-00162713.

06-JUN-2002; 2002US-00162713.

(ARCH) ARCHER-DANIELS MIDLAND CO.

Choi E, D'elia J, Kim H, Kim M, Lee JK, Pan J, Stoddard SF;
Yum D;

WPI: 2004-052025/05.
N-PSDB; AD134114, AD134122.

PT New proteins of sorbitol dehydrogenases and cytochrome c of the strains
PT Ketogulononicigenium spp., useful in molecular biology, bacteriology and
PT industrial fermentation, e.g. for producing L-sorbose and 2-keto-L-
XX gulonic acid.
XX
XX

PS Claim 67; SEQ ID NO 5; 68pp; English.

XX The invention relates to the identification and isolation of nucleic acid
CC sequences and proteins of sorbitol dehydrogenases (SDH) and cytochrome c
CC of the strains, Ketogulononicigenium sp. The proteins and nucleic acid
CC molecules are useful in the fields of molecular biology, bacteriology and
CC industrial fermentation specifically for producing L-sorbose and 2-keto-L-
CC gulonic acid (2KLG). The present sequence is Ketogulononicigenium sp. SDH
CC protein.
XX
XX

SO Sequence 578 AA;

Query Match 83.7%; Score 2569.5; DB 8; Length 578;
Best Local Similarity 80.1%; Pred. No. 7e-203;
Matches 464; Conservative 59; Mismatches 55; Indels 1; Gaps 1;

QY 1 MNPTTLRTSAVLLTAPAAFAQVPTITDELLANPPAGEMWINGRQENYRHSPLTQIT 60
DB 1 MKNSTLLASVAAVAFVAPFADVTPTDELLANPPAGEMWISYGRQENYRHSPLNQIT 60
QY 61 ADVNGQLQVWARGMEAGAVQVTPMIDGVWYLANPQDVTOALDACTGDLIWEHRQLPA 120
DB 61 PDNNGQLQVWARGMEAGAVQVTPMIDGVWYLANPQDVTOALDACTGDLIWEHRQLPE 120
QY 121 VATNAGDGRKRGVALGTSLYFSWMNHLIALDMRGVAVPVERSGEGDGLTSNTGP 180
DB 121 TSTLSSLDGRKRGVALGTVNTYFVSWDHWVALDAASGVVFDVDRGQDERV-SNNSGP 179
QY 181 IVANGVIVAGSTCOYSPYGCFTSGHDSATGEELRNHFIPOGEBGDETWNDFEARWMT 240
DB 180 IVANGVIVAGSTCOYSPYGCFTSGHDSATGEELRNHFIPOGEBGDETWNDFEARWMT 239
QY 241 GWMQGITVDPVTNLVFGYSTGVGASSETGRTGCGTLYGNTTFEAVRPTDGEIVMHHQTL 300
DB 240 GWMQGITVDPVTNLVFGYSTGVGASSETGRTGCGTLYGNTTFEAVRPTDGEIVMHHQTL 299
QY 301 PRDWMDOECTEFEMVAVNDVOPSAEMEGELAINPNAATGERRVLTGAPCKTGTWMSFDA 360
DB 300 PRDWMDOECTEFEMVAVNDVOPSAEMEGELAINPNAATGERRVLTGAPCKTGTWMSFDA 359

Qy	421	TGTYPLPLNNAACYNADQEFSAIDVNTSAYAKLAPGPNMGRIDAIDISTGRITLMSA	480
Db	420	SGTYPLPLNNAACADLAADQEFPAADVNTSAYTYLLAPEKEMGRIDAIDISTGRITLMSV	479
Qy	481	ERAAATYSPPLSTRAGGVAVNGGDRFRRLSGETSEITLMQARLAVAGQALSYELDGVQ	544
Db	480	ERLAATYSPPLSTRAGGVAVNGGSDRFRRLSEETSEITLMQTRLATVAVSGQALSYELDGVQ	538
Qy	541	YIALGAGGLTYGTQLAAPLAELADISTSVGNALTYVFALPQ	579
Db	540	YVALAGGNTYGTNLMSNIGATITDSISGNAYVFALPQ	578

Query Match	83.4%	Score 2559.5	DB 8	length 579
Best Local Similarity	80.5%	Pred. No. 4.7e-202		
Matches 466	Conservative 52	Mismatches 60	Indels 1	Gaps 1
QY	1	MNP	TLIR	SAVLLLTAPAAFOVTPITDELLANPAGEMINXGRQENYHRSPLTOT 60
DB	1	MKT	SLFPGV	ALIASYGTLLADVTPTDELLANPAGEMISIGRQENYHRSPLNQT 60
QY	61	ADNV	GOLQV	WARGMEAGAVOYTPMIDGVYVLANPBDVIQALDAQTDLIWEHRQDPA 120
DB	61	PEVNG	QOLQV	WARGMAGKAVQYTPLIHGVYVLANPDDIQALDAKTGDIWEHRQLPN 120
QY	121	VATL	NQGRKRGV	ALYGTSLYFSSMDNHLIALMERTGOVFPVVERGSGEDGLSTNTP 180
DB	121	VATL	NSFGP	PIGILALYGTINIVYFSSMDHLIALMAATGQVTFVDYDQSGED-MVSNSSGP 179
QY	181	IVANG	IVVAGST	COQSPYGCPLISGHSATGSELNRNHFIPQPBEGDTEWGNDFEARMNT 240
DB	180	IVANG	IVVAGST	COQSPGCFSGSDATGSELNRNHFIPKASBEGDETWGNDFEARMNT 239
QY	241	GVWG	QITTPY	PVNLVYFSGTGVGSPSEORCTPGCTLYGNTSPFAVSPDTGELVMAHQTL 300
DB	240	GVWG	QITTPY	PVNLVYFSGSSAVGPSPSEORCTGCTGNTSPFAVSPDTGELVMAHQTL 299
QY	301	PRDND	QECTF	PEMVYANDVOPSAEMEGRLAINPNAATGERRVLTGAPCKTGTWMSFDA 360
DB	300	PRDND	QECTF	PEMVYANDVOPSAEMDGLKISINPNAATGERRVLTGAPCKTGTWMOFDAE 359
QY	361	SGEFL	NARCTY	NTNMIASIDETGATVWEDAVLPELVEYDVCPTEFAGGRDMSALNPD 420
DB	360	TGEFL	NARCTY	QNMISIDETGATVWEDAILLDDTDTRICPTFLGGRDMSALNPD 419
QY	421	TGIV	PL	NACDIMAVDQESALDVNTSATKLAPEGEMNGRIDALDISGRILTNSA 480
DB	420	SGIV	PL	NACADIAAADOEFETADVNTSATYLLAPEKENNGRIDALDISGKLTWLV 479
QY	481	ERPA	NYSPV	LSLSTAGGVFNAGTDRYFPAISOETGELMOPARLATATVAGQALSYTELDGVO 540
DB	480	ERL	ASNYSPV	LSLSTAGGVLFNGGSDRYPALISEETGELTMQTRLATATVAGQAVSYELDGO 539
QY	541	YIAI	AGGLT	YGTQLMAPLAELDSTSGNALIYFALPO 579
DB	540	YIAI	AGGLT	YAVONRPLAEVSTSGNALIYFALPO 578

PR 19-SEP-1996; 96EP-00115001.
 XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX
 PI Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
 XX MPI; 1998-195228/18.
 DR N-PSDB; AAV29052.
 XX
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde dehydrogenase
 PT enzyme(s) - useful for converting substrate(s), e.g. L-sorbose or D-
 PT sorbitol to 2-keto-L-gulonic acid.
 XX
 PS Claim 1; Page 38-40; 59pp; English.
 XX
 CC This is the amino acid sequence for the Gluconobacter oxydans alcohol
 CC and/or aldehyde dehydrogenase A' enzyme. The enzymes or recombinant
 CC organisms can be used to convert suitable substrates to aldehydes,
 CC ketones or carboxylic acids, especially to convert L-sorbose or D-
 CC sorbitol to 2-keto-L-gulonic acid, which can be converted to L-ascorbic
 CC acid by standard procedures. The derivatives of ADH enzymes have desired
 CC substrate specificity, higher affinity to a substrate, lower affinity to
 CC an inhibitory compound, higher stability against temperature and/or pH
 CC and higher catalytic speed
 XX
 SQ Sequence 579 AA;
 Query Match 82.3%; Score 2526.5; DB 2; Length 579;
 Best Local Similarity 79.8%; Pred. No. 2.5e-199;
 Matches 462; Conservative 56; Mismatches 60; Indels 1; Gaps 1;

1 MNPTTLRTSAVLLTAPAFQVTPITDELLANPAGEMINYGNOENYRHSPLTQIT 60
 1 MKTSSILVAVSALASVSSFALQVTPVDELLANPAGEMISYGNQENYRHSPLTQIT 60
 QY 61 ADVNGQLQVWARGMEKAVQVTPMTHDGVMTLANPQDVIOALDACTGLIMEHRQLPA 120
 DB 61 TENVGQLQVWARGMEKAVQVTPMTHDGVMTLANPQDVIOALDACTGLIMEHRQLPN 120
 QY 121 VATLNQGRKRGVALYGTSLYFSSMDNHLIALDMETGVVFPVVERSGSDGLTSTTGP 180
 DB 121 IATLNSGEPTRGMALYGTINVTFSMDNHLIALDTATGQVFPVDRQGTGSD -MVSNSGP 179
 QY 181 IVANGVIVAGSTQCYSPFGCFISGHSATGEBELMRNHFIPQEGEGDETWGDFEARWMT 240
 DB 180 IVANGVIVAGSTQCYSPFGCFVSGHDSATGEBELMRNHFIPRAGEEGDETWGDFEARWMT 239
 QY 241 GVNQGITVDPVTMLVFSGTGVGPASSETQRTGRTGLYGNTRFAPRPTGRIVWRHQT 300
 DB 240 GVNQGITVDPVTMLVFSGTGVGPASSETQRTGRTGLYGNTRFAPRPTGRIVWRHQT 299
 QY 301 PRDNDQECTFEMNVANVDVQPSAEMEGLRAINPNATGERVLTGAPCKTGIMWSFDA 360
 DB 300 PRDNDQECTFEMNVANVDVQPSADMDGVRSINPNATGERVLTGVPCKTGIMWQFDAE 359
 QY 361 SGEFIMARDNTYMTLASIDETGLVYNEBAVYKEIDVEYDVPTFLGSRDSSALAND 420
 DB 360 TGFELVARDTSYENILLESIDENGIYVDESKVLTEDTPVDCVFLLRDWPSSALAND 419
 QY 421 TGEYFLFNACVDINAVDQESFALDVYNTATAKIAPGFEMNGRIDALIDISTGRTIWSA 480
 DB 420 TGEYFLFNANTCMDEIVDQESFSLDVYNTSLAKAPKELVGRIDALIDISTGRTIWSA 479
 QY 481 ERPAANTSPLYLSTAGGVNVTGTRIFRALISQETGTLWQARLATYATQOASIELDQV 540
 DB 480 ERESNANVAPLSTAGGVNVTGTRIFRALISQETGTLWQARLATYATQOASIELDQV 539
 QY 541 YIAIAGAGLTGTQOLNAPLAEADISTSVGNAYVFPALPQ 579
 DB 540 YIAIGGGGTTTGGSFHNRPLAEPVDSTALIGNAMVFPALPQ 578

RESULT 8

AAW37875
 ID AAW37875 standard; protein; 578 AA.
 XX
 AC AAW37875;
 XX
 DT 10-AUG-1998 (first entry)
 XX
 DE Alcohol and/or aldehyde dehydrogenase A' amino acid sequence.
 XX
 KW Alcohol/aldehyde dehydrogenase A' enzyme; recombinant organism;
 KW aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
 KW 2-keto-L-gulonic acid; L-ascorbic; inhibition.
 XX
 OS Gluconobacter oxydans.
 XX
 XX Key Location/Qualifiers
 XX FH Peptide 1..23
 XX FT Protein /note= "signal peptide"
 XX FT Protein 24..578
 XX FT Protein /note= "mature protein"
 XX
 XX EP832974-A2.
 XX
 XX 01-APR-1998.
 XX
 XX 11-SEP-1997; 97EP-00115801.
 XX
 XX 19-SEP-1996; 96EP-00115001.
 XX
 XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX
 XX Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
 XX MPI; 1998-195228/18.
 XX
 XX N-PSDB; AAV29053.
 XX
 XX Recombinant Gluconobacter oxydans alcohol and/or aldehyde dehydrogenase
 XX enzyme(s) - useful for converting substrate(s), e.g. L-sorbose or D-
 XX sorbitol to 2-keto-L-gulonic acid.
 XX
 XX Claim 1; Page 41-43; 59pp; English.
 XX
 CC This is the amino acid sequence for the Gluconobacter oxydans alcohol
 CC and/or aldehyde dehydrogenase A' enzyme. The enzymes or recombinant
 CC organisms can be used to convert suitable substrates to aldehydes,
 CC ketones or carboxylic acids, especially to convert L-sorbose or D-
 CC sorbitol to 2-keto-L-gulonic acid, which can be converted to L-ascorbic
 CC acid by standard procedures. The derivatives of ADH enzymes have desired
 CC substrate specificity, higher affinity to a substrate, lower affinity to
 CC an inhibitory compound, higher stability against temperature and/or pH
 CC and higher catalytic speed
 CC
 SQ Sequence 578 AA;
 Query Match 81.4%; Score 2497; DB 2; Length 578;
 Best Local Similarity 79.3%; Pred. No. 6.7e-197;
 Matches 459; Conservative 57; Mismatches 61; Indels 2; Gaps 2;

1 MNPTTLRTSAVLLTAPAFQVTPITDELLANPAGEMINYGNOENYRHSPLTQIT 60
 1 MKTSSILVAVSALASVSSFALQVTPVDELLANPAGEMISYGNQENYRHSPLTQIT 59
 QY 61 ADVNGQLQVWARGMEKAVQVTPMTHDGVMTLANPQDVIOALDACTGLIMEHRQLPA 120
 DB 60 ADVNGQLQVWARGMEKAVQVTPMTHDGVMTLANPQDVIOALDACTGLIMEHRQLPN 119
 QY 121 VATLNQGRKRGVALYGTSLYFSSMDNHLIALDMETGVVFPVVERSGSDGLTSTTGP 180
 DB 120 IATLNSGEPTRGMALYGTINVTFSMDNHLIALDTATGQVFPVDRQGTGSD -MVSNSGP 178
 QY 181 IVANGVIVAGSTQCYSPFGCFISGHSATGEBELMRNHFIPQEGEGDETWGDFEARWMT 240
 DB 179 IVANGVIVAGSTQCYSPFGCFVSGHDSATGEBELMRNHFIPRAGEEGDETWGDFEARWMT 238

QY 241 GWMGQITYPVNLVFGSTGVGPASFTQRTGPGCLYGTNTFPAPVPTGELIWMHQTL 300
 Db 239 GWMGQITYPVNLVFGSTGVGPASFTQRTGPGCLYGTNTFPAPVPTGELIWMHQTL 298
 QY 301 PRDWDCECFEEMVNAVVDQPSAEMEGRLAIPNNAATGERRLVLTGA PCKTGTMSPDA 360
 Db 299 PRDWDCECFEEMVNAVVDQPSAEMEGRLAIPNNAATGERRLVLTGA PCKTGTMSPDA 358
 QY 361 SGEFLMARDNTYNTMIAISIDETGLVTVNEDAVLKELDVBYDVCPTFGSDMSAALNDP 420
 Db 359 TGEFLMARDNTYNTMIAISIDETGLVTVNEDAVLKELDVBYDVCPTFGSDMSAALNDP 418
 QY 421 TGIYFLPNNACDIMAVIDEFSALDYNTSATRAKLA PGEENGRIDALIDISTGRLMSA 480
 Db 419 TGIYFLPNNACDIMAVIDEFSALDYNTSATRAKLA PGEENGRIDALIDISTGRLMSA 478
 QY 481 ERPAANTSPVLTAGVVFNGGTDRTYFRALSOETGETLMQARLATVATGQAI SYELDGVQ 540
 Db 479 ERVANSVAPVLTAGVVFNGGTDRTYFRALSOETGETLMQARLATVATGQAI SYELDGVQ 538
 QY 541 YVIAAGGLTYGTQLNAPLAELDSTISVGNALIVFALPQ 579
 Db 539 YVIAAGGLTYGTQLNAPLAELDSTISVGNALIVFALPQ 577

RESULT 9

ADBE94131
 ID ADBE94131 standard; protein; 608 AA.

AC ADBE94131;

DT 12-FEB-2004 (first entry)

DE Alcohol/aldehyde dehydrogenase, SEQ ID 1.

KM Enzyme; alcohol/aldehyde dehydrogenase; pyrrolo quinoline quinone;

XX L-sorbose; L-sorbose; 2-keto-L-gulonic acid.

OS Pseudoglucobacter saccharoketogenes; IF014464.

XX JF2003159079-A.

XX 03-JUN-2003.

XX 29-NOV-2001; 2001JP-00364508.

XX 29-NOV-2001; 2001JP-00364508.

XX (FUJI) FUJISAWA PHARM CO LTD.

XX WPI; 2003-818681/77.

XX N-PSDB; ADBE94132.

PT Novel alcohol/aldehyde dehydrogenase protein catalyzing oxidation of
 PT hydroxymethyl group of compound to aldehyde group and aldehyde group of
 PT compound to carboxyl group, useful for manufacturing 2-keto-L-gulonic
 acid.

PS Claim 3; SEQ ID NO 1; 48pp; Japanese.

CC The present invention relates to an alcohol/aldehyde dehydrogenase (I;
 CC ADBE94131), which catalyzes oxidation of hydroxymethyl group of a compound
 CC to an aldehyde group and the aldehyde group of a compound to a carboxyl
 CC group. (I) does not contain hemoferrum and rare earth elements. (I) has
 CC an optimum pH of 4.5-5.5, isoelectric point of 4.1 +/- 0.3, pyrrolo
 CC quinoline quinone as a prosthetic group and K_m value of 40 mM for
 CC sorbose. (I) is useful for manufacturing a compound having a carboxyl
 CC group by contacting the compound having hydroxymethyl group or an
 CC aldehyde group with (I) where the compound having a hydroxymethyl group
 CC or an aldehyde group is L-sorbose or L-sorbose and the compound having
 CC a carboxyl group is 2-keto-L-gulonic acid.

SQ Sequence 608 AA;

Query Match 54.5%; Score 1673; DB 7; Length 608;
 Best Local Similarity 52.5%; Pred. No. 6.9e-129;
 Matches 315; Conservative 94; Mismatches 167; Indels 24; Gaps 6;

QY 2 NPTTLIRTSANVLLTAAPAPQ-----VPIITBELLANPAGWI 42
 Db 9 NVVGLATSTALIASISGP-AFAQHDANAAPSKAGQSAIENFQVTVTDLDLAKXNPANMP 67
 QY 43 NYGRQENVRHSPPLTQITADNVGCOLVWARGMEAGAVQTPMIDGMYLANPQDVIOA 102
 Db 68 ILRGNYQGMGVSPDLQIKNDVNGDLQVMSRTMEGSGAIAANGVIFIGNTNDVIOA 127
 QY 103 LDAQTGDLIMEHRROLPAVAT-LNAQGRKGVALYGTSLYSSMDNLILDMETGVV 161
 Db 128 IDGKTGSLIMEYRRRLPSASKFINSILGAAKSIALFGDKVYFVSVDNFFVALDAKTXGLA 187
 QY 162 PDVERGSGEDGLTNTGTPTIVANGVIVAGSTCOXSPYGCFTSGHDSATGEELMRNHFIPQ 221
 Db 188 WETNRGQYERBVAUSSGPVIVDGVILAGSTCOXSGFCYVGTDAESGEELMRNHFIPR 247
 QY 222 PGEEDGETWGN-DFEARMWTGVWQITVPTNLVFGSTGVGPASFTQRTGPGCLYGT 280
 Db 248 PGEEDDITWCGAPYENRMWTGAMQITVDPEDILVYSGTGAGPASEVQRTGEGTLAGT 307
 QY 281 NTRFAVRPDTGELIWMHQTLPRDNDQCTFEMVNAVVDQPSAEMEGRLAIPNNAAGE 340
 Db 308 NTRFAVRPDTGELIWMHQTLPRDNDQCTFEMVNAVVDQPSAEMEGRLAIPNNAAGE 367
 QY 341 -RRVLTGAPCKTGTMSPDAASGEFLMARDNTYNTMIAISIDETGLVTVNEDAVLKELDVE 399
 Db 368 TEKVLTVGPCKTGVMQDPAKTGDFWEGKATVEQNSIASIDTGLVTVNEDMILKEPKXT 427
 QY 400 YDVCPTFGGRDMSAALNPDTGTYFLPNNACDIMAVIDEFSALDYNTSATRAKLA PG 459
 Db 428 YNYCPTFLGGRDMSAALNPDTGTYFLPNNACDIMAVIDEFSALDYNTSATRAKLA PG 487
 QY 460 FENMGRIDALIDISTRTLSAERPAANTSPVLTAGVVFNGGTDRTYFRALSOETGETLM 519
 Db 488 KTNMGRVADILATSETKMSYETRALYDPVLTGGLDVFVGGIDRDRAIDBSGKGVW 547
 QY 520 QARLATVATGQAI SYELDGVYIAGAGLTYGTQLNAPLAELDSTISVGNALIVFALPQ 579
 Db 548 STRLPGAIVSGYTTSYIDGRQYAVVSGG-SLGGPTFGPTTDPDVASGANGIYVFLPQE 606

RESULT 10

AAB35987
 ID AAB35987 standard; protein; 754 AA.

AC AAB35987;

DT 01-MAR-2001 (first entry)

DE Sorbitol dehydrogenase subunit 1 amino acid sequence.

KM Sorbitol dehydrogenase; SDH; Glucobacter oxydans; D-sorbitol;

XX L-sorbose production; 2-keto-L-gulonic acid.

OS Glucobacter oxydans.

XX WO200065066-A1.

XX 02-NOV-2000.

XX 23-APR-1999; 99WO-1B000736.

XX 23-APR-1999; 99WO-1B000736.

XX (CHOI/) CHOI E.

XX (RHEB/) RHEB S.

XX (LEEB/) LEE E.

XX Choi E, Rhee S, Lee E;
 XX WPI: 2000-687351/67.
 DR N-PSDB; AAC83153.
 XX
 PT Novel membrane-bound sorbitol dehydrogenase nucleic acid molecule
 PT isolated from Gluconobacter suboxydans useful for the fermentative
 PT production of 2-keto-L-gulonic acid and L-sorbose from D-sorbitol.
 XX
 XX Claim 1; Fig 8; 96pp; English.
 CC This invention relates to an isolated membrane-bound sorbitol
 CC dehydrogenase (SDH) from Gluconobacter oxydans. The invention includes
 CC polynucleotides (AAC83153 - AAC83154) encoding the 3 subunits of SDH
 CC (AAB35987 - AAB35989). Also included in the invention are two
 CC polynucleotide sequences AAC83156 and AAC83157 which contain the subunit
 CC coding regions. The sorbitol dehydrogenase polynucleotide sequences are
 CC useful for producing L-sorbose from D-sorbitol and for increasing the
 CC production of 2-keto-L-gulonic acid by transforming a host cell,
 CC especially Gluconobacter with the DNA and selecting the transformed host
 CC cell. Note: An additional 66 DNA sequences (fragments of the SDH subunit
 CC genes) encoding fragments of SDH are specifically claimed, however these
 CC sequences are not given separately in the specification but are included
 CC in sequences AAC83156 and AAC83157
 CC
 XX Sequence 754 AA;
 SQ
 Query Match 19.1%; Score 585; DB 3; Length 754;
 Best Local Similarity 28.2%; Pred. No. 5,5e-39;
 Matches 172; Conservative 94; Mismatches 238; Indels 106; Gaps 21;
 6 LRTSAAY-LLTAPAFAG--VTPTDELANPAGMINTGRNENYRHSPLDTAD 62
 16 LIGCAALAFCATSPALMEDTGTALTN--ADQHPGDMWSYKRTTSEORISFLDITD 72
 63 NVGQLOLVW-----ARGMEAGAVOYTPMIDGVYLANPGVIOALDQGLDWEHR 116
 73 NASNLKLAHYDDITNRQGE-----TPLIDGVMTATNWSMKALDADATKLTMSYD 127
 117 QLPRA-VATLNAQGRKRGVALYGTSLYFSSNDHLALDMENQGVYFV----- 164
 128 KVPENIADRCCTVNRGAAYWNGKYFETPDRLLALAKTKLWASVYVTKRQOLH 187
 165 ERSGEDELSTNTGPIVANG-VIVAGSTCOYSPYCGFISGHDSATGEELMRHFIPCP 223
 188 QRSYTDG-----APRIAKGXIIIONGAEFGARG-FTYAVDAETGKMDMRFETPNPD 240
 224 EEGD-----ETWQNDPEARMTG--VWGQITPYNTLVYFGSTGVGPASE 267
 241 NKPDGAASDVLMSKAYPTWKGGAWKQGGGGTVDLSLYPVTDLVYLGVNGSPMY 300
 268 TQGTGPGTLYGNTFRPVRPDTGETVYRHTLPRDNMQOECTFEEMVANVDVQPSAEWE 327
 301 KFSSEGKNNLFLGSIVAINPDTGKYVMHGETPMQMDYTSQCLMA--LDMPVNGEN- 357
 328 GLRAINPNATGERRVLTGAPCKTGTWMSFDAASGEFLMARDNTYNTMIASIDE-TGLVT 386
 358 -----RHVTVHAP-KNGFFYIIDAKTGKISGKPYTYENMANGLDPVGRPN 403
 387 VNEDAVYLKEDVYDVOTFLGGRDWSAALNDPTGIYFLPLNNACVYDIMAVDQESAI 446
 404 YNPDALMTLNGKRWYGPDLGSHNPAVAAYSPQTKLVYIPAOQVF--VTDPOGKXK 460
 447 VYNTSATKLAPGFENGRIDAIDISTGR-----TLMSAERPAANY-- 488
 461 AHHDSNMLGL-----DNKRIGLDDNDPQHKADKAQFLKDLKGMIVAMPQKQAQAFVD 515
 489 -----PVLTAGGVFNNGTDRYPRALSQETGETLMQARLATATGALISYELDGYX 541
 516 HKSPMNGGLLATAGGVLFQGLANGEPHAYDATTGKDLFTTPAOSATIAAPVITYANGX 575
 542 IA--IGAGGL 549

DB 576 VAVEWGMGI 585
 RESULT 11
 ID ABO83287
 XX ABO83287
 AC ABO83287;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Pseudomonas aeruginosa polypeptide #15462.
 XX
 KM Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
 XX
 OS Pseudomonas aeruginosa.
 XX
 PN US651795-B1.
 XX
 PD 22-APR-2003.
 XX
 PF 18-FEB-1999; 99US-00252991.
 XX
 PR 18-FEB-1998; 98US-0074788P.
 XX
 PR 27-JUL-1998; 98US-0094190P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Rubenfield MJ, Nollig J, Deloughery C, Bush D;
 XX
 DR WPI: 2003-615309/58.
 XX
 DR N-PSDB; ABD16858.
 XX
 PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.
 XX
 PS Disclosure; SEQ ID NO 32033; 455pp; English.
 XX
 CC The invention relates to Pseudomonas aeruginosa polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics,
 CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-P. aeruginosa drugs, as templates for recombinant
 CC production of P. aeruginosa-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of P. aeruginosa-caused
 CC infection, and in detection of P. aeruginosa sequences or other sequences
 CC of Pseudomonas species using biochip technology. Sequences ABO67825-
 CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html
 CC
 XX Sequence 685 AA;
 SQ
 Query Match 18.0%; Score 552; DB 7; Length 685;
 Best Local Similarity 27.2%; Pred. No. 2,5e-36;
 Matches 180; Conservative 110; Mismatches 220; Indels 142; Gaps 28;
 3 PTTLTISA-----AVLLTPPAFAQVTPITDELAN--PPAGMINTGRNENYRHS 55
 71 FAGLRPSLHCLAPRAVALGSGAALAK--DVTMEDIANDDKTTGDVLYGMSGTHAQKWS 128
 56 LTQTTADVQLOLVMA-----RGEAGAVOYTPMIDGVYLANPGDVIOALDAQT 107
 129 LKQVNAADVFLCTRAVSYSGDEKRGQESQAT-----VSDGVITYVAYSRLFALDAKT 163
 108 GDLIWEHRRL-----PAVATLNAQGRKRGVALYGTSLYFSSNDHLALDMETGVYF 162

Db 184 GKRLMTYNHRLPDDIRECCDVN-----RGAALYGDKVFFGLDASVALNKNTKYVW 237
 QY 163 DVERSGEDGLTSNTTGPVANG-----VIVAGST-CQYSPYGCFTSGHDSATGEIWMEN 216
 Db 238 KKRPADHGAATYMTGATITVKGDKTGKVLIIHSSSGRFRVRLRA-RPDPGEIWMR 296
 QY 217 HFLPQPEEG-----DETGNDEPA-----RMWTG---WGQITYPVTN 253
 Db 297 PFV--EGHMGRLNGKSTVTGDKVAPSPDDRNSPTGKVESWGHGGAPQASFAETN 354
 QY 254 LVFVSTGVGPASETORGPFG-----TLVGTNTPRAVPTDGEIYWRHQLPRPNWQ 307
 Db 355 TIVVGAANPFPMTWMTATAGKGNPHYDSLX-TSGQGVDPSSGEVWMTFYQHPNDAMP 413
 QY 308 ECTFEMVNAVVDVQPSAEMEGELRAINPNATGERPVLTPGPKTGTMSEDAAGE----- 363
 Db 414 SGNNEVL-----FDYKADGKIVKATAHADNNGFFYVDRNGKLQNA 457
 QY 364 -----FLWADNTYTMIMASID-ETGLVTVNED--AVLKEIDVEY---DVCEPFLGR 410
 Db 458 PPFVDNTTWA-----SHIDIKTRPVREGRQRPLEPQCKGKAVEVSPFLGK 508
 QY 411 DWSSAALNPDTGTYFLPLNNAACYDIMAVDQESALDVYNTSATKLAPEF-----ENK 463
 Db 509 NMPNMAVSQDTGLFYVANH-----WKEDVTEBSYTKGSAVILGMGRIRKMDY 561
 QY 464 GRIDAIISTGRILMSAERPANYSPLYSTAGGVENGSTDRYFRALSOETGETLMQAL 523
 Db 562 GSRFAMPVSGKVWMEKEHLPLMAGVLTATAGNLVFTGDDGYFKAFDAKSGKELWKFOT 621
 QY 524 ATATGGAISVELDGVQY--IAIGAGGLT--YG--TQNLAPLAELIDSTSVGNALYVA 576
 Db 622 GSGEVSPTIWEQDEQYLGVTGYGSAVPLMGDMADLTRPVAQ-----GSGTWVK 674
 QY 577 LP 578
 Db 675 LP 676
 RESULT 12
 AAR20192 standard; protein; 738 AA.
 XX AAR20192;
 AC AAR20192;
 XX 27-AUG-2003 (revised)
 DT 16-APR-1992 (first entry)
 XX ADH complex protein (mol.wt. 72.000).
 DE
 XX Alcohol dehydrogenase; acetic acid; fermentation.
 KM
 XX Acetobacter sp.
 OS
 XX JPO3266988-A.
 PN
 XX 27-NOV-1991.
 PD
 XX 26-MAR-1990; 90JP-00073440.
 PE
 XX 26-FEB-1990; 90JP-00042391.
 PR
 XX (NAKA-) NAKANO SUTEN KK.
 PA
 XX Tamaki T, Takemura H, Tayama K, Fukaya M, Okumura H, Kawamura Y,
 PI
 XX WPI: 1991-288462/40.
 DR
 XX N-PSDB; AAQ20383.
 DT
 XX Gene for membrane-bound alcohol dehydrogenase complex - obcd, from
 PT
 XX Acetobacter alioacetigenes, used for prodn. of enzyme for converting
 PT
 XX alcohol to acid.
 XX

PS Disclosure; Fig 3 (1-3); 21pp; Japanese.
 XX
 CC Acetobacter transformed with the sequence encoding this protein can
 CC enhance the efficiency of acetic acid fermentation. The ADH complex can
 CC be easily extracted from the bacteria and purified and it can be used for
 CC the determination of an alcohol. See also AAQ20383-84, and -86-88.
 CC (updated on 27-AUG-2003 to correct OS field.)
 CC
 SQ Sequence 738 AA;
 Query Match 18.0%; Score 551.5; DB 2; Length 738;
 Best Local Similarity 27.0%; Pred. No. 3,1e-36;
 Matches 164; Conservative 91; Mismatches 253; Indels 99; Gaps 17;
 QY 5 TLRTSAVLLLTAPAAEQVTP--ITDELL--ANPPAGEMINYGENQENYRHSPLTOI 59
 Db 15 TLTAGTICAAALISGATWASADGGGATGEALIHADPHGNMWTYGRFYSQRYSPLDQI 74
 QY 60 TADNVGQQLVW-----ARGMEAGAVQVTEPHIDGVVYLANPBDVIALDAQGDLWE 113
 Db 75 NRSNVGNLKLAWYLDLDTNRQEG---TLVVDGVVYATTNMSMKVAADAAATGLMS 129
 QY 114 HRRQLPA-VATLNAQGRKRGVALYGTSLYFSSMDNHIALDMETGVYFV----- 164
 Db 130 YDPRVPGNIADKCCDDTVNRGAAYVNGKVYFGTFGRILALDAKTGKLVMSVNTIPEAE 189
 QY 165 ---ERSGEDGLTSNTTGPVANG-VIVAGSTCQYSPYGCFTSGHDSATGEIWMRHFIP 220
 Db 190 LKQKSYTVDG-----APRIAKGRVITIGNGSGSEFGANG-FVSAFDAETGVDRFFTVP 242
 QY 221 QPGEED-----ETMGNDFEARMWTG-----VWGQITYPVTNLVPEV 258
 Db 243 NPKMEPDASDVLNKAIVQTS-----PTGANTRGGGGGTWDSLVIPVADLVYLG 295
 QY 259 STGVGPASETORGPFGTLVGTNTRFAVRPDTGEIYWRHQLPRPNMDQECTFEMNAV 318
 Db 296 VNGSGPMVYKYRESBEGGDNLEFGSIVALKPTEGVYWFQETPMQWDFTSDDQIMTIDL 355
 QY 319 DVQPSAEMEGELRAINPNATGERPVLTPGPKTGTMSEDAAGEFLWARTNTNMTAS 378
 Db 356 PI-----NGSTRVIVHARKNGFFYIIDKKTGEFISGKNYVYVWNASG 398
 QY 379 ID-ETGLVTVNEDAVLKEIDVEYVCPTFLGGRDMSALNPDTGIYFLPLNNAACY--- 433
 Db 399 LDPKTGRPIYNDALYTLTGKEMWGIPEGLGSHNPAANA-FPKGLVYIIPAOVPELYTN 458
 QY 434 -----DIMAVDQESALDVYNTSATKLAPEFEMGRIDALIDISTGRILMSAERPA 484
 Db 459 QVGFPTPHDSDWNLGLDMNKVGIPID-SPEAKQAFVKDLKGIIVAWDPQKQAEARVHDKG 517
 QY 485 ANYSPVLSTAGGVENGSTDRYFRALSOETGETLMQALATVATGQALSYELDGVQYLA- 543
 Db 518 PWNQGIATGGDLFQGLANGEFHAYDATNGSDLFHFAADSGIILAPRTYILANGKQYAV 577
 QY 544 -IGAGGL 549
 Db 578 EVWGMDGI 584
 RESULT 13
 AAR13993 standard; protein; 738 AA.
 XX AAR13993;
 AC AAR13993;
 XX 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 09-DEC-1991 (first entry)
 XX A.alioacetigenes membrane-bound ADH 72kD sub-unit.
 DE
 XX alcohol dehydrogenase complex; carboxylic acid production.
 KM
 XX

```

OS Acetobacter sp.
XX EP448969-A.
XX
XX 02-OCT-1991.
XX
XX 26-FEB-1991; 91EP-00102793.
XX
XX 26-FEB-1990; 90JP-00042391.
XX
XX 26-MAR-1990; 90JP-00073440.
XX
XX (NAKA-) NAKANO VINEGAR CO LTD.
XX
XX Tamaki T, Takemura H, Tayama K, Fukaya M, Okumura H, Kawamura Y;
XX WPI; 1991-289462/40.
XX
XX N-PSDB; AAQ13580.
XX
XX Gene for membrane-bound alcohol dehydrogenase complex - obtd. from
XX Acetobacter alcohigenes, used for prodn. of enzyme for converting
XX alcohol to acid.
XX
XX Disclosure; Fig 3; 36pp; English.
XX
XX Total DNA was prepared from A.alcohigenes MH-24, digested with PstI and
XX SmaI and ligated to PstI-SmaI cleaved pUC18. The ligation mixture was
XX used to transform E.coli JM109. Probes were designed based on the N-
XX terminal amino acid sequence of the ADH complex isolated from A.
XX alcohigenes (see AAQ13582-Q13584). The gene encoding the 72kD ADH
XX subunit was isolated and sequenced. The directly sequenced N-terminal
XX region of the purified 72kD sub-unit corresponds to the sequence
XX beginning at residue 36 of the deduced sequence. This suggests that the
XX first 35 N-terminal amino acids form a leader peptide involved in
XX secretion of the mature 72kD protein. The deduced amino acid sequence has
XX 77 percent homology with the same enzyme from A. aceti K6033. See also
XX AAQ13581. (Updated on 25-MAR-2003 to correct PA field.) (updated on 27-
XX AUG-2003 to correct OS field.)
XX
XX Sequence 738 AA:
XX
XX Query Match 17.9%; Score 550.5; DB 2; Length 738;
XX Best Local Similarity 27.0%; Pred. No. 3.7e-36;
XX Matches 164; Conservative 91; Mismatches 253; Indels 99; Gaps 17;
XX
XX 5 TLRTSAVLLTAPAFQVTP--ITDELL--ANPAGEMINRGNOENRHSPLTQI 59
XX ||| : : : : : : : : : : : : : : : : : : : : : : : : :
XX 15 TLTAGTICALLISGYATMASADGQATGEAIIHADHPGNMTYRTYSDDRYSPDQI 74
XX ||| : : : : : : : : : : : : : : : : : : : : : : : : :
XX 60 TADNVGQLQVW-----AGMEAGAVQVTPMTHDGMVTLANGDVITQALDQGLDLYNE 113
XX ||| : : : : : : : : : : : : : : : : : : : : : : : : :
XX 75 NRSNVGMLKLAWLDLDTNRGQEG--TPVLIDGVMTATTNWSKAVDAATGKLWS 129
XX ||| : : : : : : : : : : : : : : : : : : : : : : : : :
XX 114 HRROLPA-VATLNAQGRKRGVALYGTSLYSSMDNHLALDMETQGVFV----- 164
XX ||| : : : : : : : : : : : : : : : : : : : : : : : : :
XX 130 YDGRVFNINADKGGCDVNRGAAYMNGKYVFGFFDRRLALDAKTLKWSVNTIPPEAB 189
XX ||| : : : : : : : : : : : : : : : : : : : : : : : : :
XX 165 ---ERGGEGGLTSNTTGPVANG-VIVAGSTQCYSPYCGFISGHSAAGEELMRHFRP 220
XX ||| : : : : : : : : : : : : : : : : : : : : : : : : :
XX 190 LGRQRYSTTVVG-----APRIAGRYLIIGNGSEFARG-FVSAFPAELGKXVWREFTPV 242
XX ||| : : : : : : : : : : : : : : : : : : : : : : : : :
XX 221 QPGEEDP-----ETWGNDFEARMTG-----VMQITVDPTVNTLVFG 258
XX ||| : : : : : : : : : : : : : : : : : : : : : : : : :
XX 243 NPKNEPRAADSVIANKAYQVNS-----PTGAMTRQGGCGGVMSIYDPAADLVYIG 295
XX ||| : : : : : : : : : : : : : : : : : : : : : : : : :
XX 259 STGVGPASETQGTGGTLYGNTREPAVRDPDGEIYVRHQTLEFRDWDQECFEMNVANV 318
XX ||| : : : : : : : : : : : : : : : : : : : : : : : : :
XX 296 VGNQSPNNYKRSRGKDNLFILGIVALKPEDEYVWHFQETPMQDMDFTSDDQINTLIDL 355
XX ||| : : : : : : : : : : : : : : : : : : : : : : : : :
XX 319 DVQPSAMEGLRALINPAAATGERRVLTGAPCKTGTMWSPFAASGEFLMRADNTYNTMAS 378
XX ||| : : : : : : : : : : : : : : : : : : : : : : : : :
XX 356 PI-----NGETRHVIYHARKNGFFIYIDANTGEISGKNYYVYVMASG 398
XX ||| : : : : : : : : : : : : : : : : : : : : : : : : :
XX 379 ID-ETGLVTNEDAVLKEIDVEYDVCFPLGGRDWSAALNPDITGIYPLNNACY--- 433
XX ||| : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

DB 399 LDPKTRPIYDPLLYLTITGKEMVIGIPGDLGHNFAAMAFSPKGLVYITPAQGVPEFYTN 458
XX ||| : : : : : : : : : : : : : : : : : : : : : : : : :
XX 434 -----DIYAVDOEFSLADVNTSATAKLAPGFENMRIDAIDISTRTIWSAERPA 484
XX ||| : : : : : : : : : : : : : : : : : : : : : : : : :
XX 459 QVGFTPHDPDSMNLGDMNRKGVIPD-SPEAKQAFVVKDKMWIVAWDPQKQAEAMRVVHK 517
XX ||| : : : : : : : : : : : : : : : : : : : : : : : : :
XX 485 ANYSPVLTAGGVVFNNGTDRYFPALSOFTGETLMQARLATAVNTGQALSYELDGVYIA- 543
XX ||| : : : : : : : : : : : : : : : : : : : : : : : : :
XX 518 FNGGILATGDDLFPQGLANGEFHAYDATNGSDFHFAADSGIITAPVYTLANGKQYVAV 577
XX ||| : : : : : : : : : : : : : : : : : : : : : : : : :
XX 544 -IGAGGL 549
XX ||| : : : : : : : : : : : : : : : : : : : : : : : : :
XX 578 EVWGGI 584
XX
XX RESULT 14
XX AAR05235
XX AAR05235 standard; protein, 742 AA.
XX
XX AAR05235;
XX
XX 24-OCT-2003 (revised)
XX 25-MAR-2003 (revised)
XX 04-AUG-1990 (first entry)
XX
XX Amino acid sequence of alcohol dehydrogenase (ADH) encoded by bases 1-
XX 2229 of alcohol dehydrogenase (ADH) gene.
XX
XX Alcohol dehydrogenase (ADH) gene; Acetobacter pasteurianus IFO 3191;
XX Acetobacter aceti K1006 (FERM-7528); plasmid PAA121.
XX
XX Acetobacter aceti; strain K1006 (FERM-7528).
XX
XX JP02000452-A.
XX
XX 05-JAN-1990.
XX
XX 29-MAR-1988; 88JP-00075069.
XX
XX 30-OCT-1987; 87JP-00273190.
XX
XX (QPP) OP CORP.
XX (QPP) KEMPIE JOZO KK.
XX
XX WPI; 1990-047990/07.
XX N-PEDB; AAQ91811.
XX
XX Cloning DNA, plasmid and microbe contg. it - contains alcohol
XX dehydrogenase gene derived from acetobacter aceti K1006 strain.
XX
XX Disclosure; Fig 4; 8pp; Japanese.
XX
XX Also new are a recombinant plasmid contg. its encoding DNA, and a
XX microorganism transformed with the plasmid. A DNA fragment was obtd. from
XX A. aceti strain K1006 and was transferred to an ADH-defective strain. The
XX resulting plasmid, PAA121, contg. its gene was inserted into A.
XX pasteurianus strain IFO 3191 by the triparent method. Its gene is useful
XX for improving Acetobacter culturing. (Updated on 25-MAR-2003 to correct
XX PF field.) (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-
XX MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS
XX field)
XX
XX Sequence 742 AA:
XX
XX Query Match 16.0%; Score 490.5; DB 2; Length 742;
XX Best Local Similarity 25.7%; Pred. No. 3.3e-31;
XX Matches 152; Conservative 102; Mismatches 256; Indels 81; Gaps 17;
XX
XX 11 AAVLLTAPAFQVTPITDELL--ANPAGEMINRGNOENRHSPLTQITADNVGQLQ 68
XX ||| : : : : : : : : : : : : : : : : : : : : : : : : :
XX 23 AALPFAVPAARAAGQNTGEAIIHADHPENMLSYGRYSRQYSPDQINSNVGDLK 82
XX ||| : : : : : : : : : : : : : : : : : : : : : : : : :

```

QY 69 LV-----WARGMAGAVQVTPMINDGWTLANPBDVITQALDAQDGLIWEHRQLPA-V 121
 PS 83 LLEGYTLDTNRGGBA-----TPLVVDGIMVATTWMSMEALDAATGLIMQDYPKPGKI 137
 Db 122 ATLNAQDGRKRGVALYGTSLYFSSWMDNHLALDMETGQVVEPDRSGSGDGLSNITGP 181
 QY 138 ADKGCCTVNRGAGYNNKGYFWGTFFDGRVLAADAKTKKMAVNTTPADSLKQKRSYV 197
 Db 182 ----VANGLVAGS--TCQYSPYGCFTISGHDATGSELMWNEHFIPOGSEBD----- 227
 QY 198 DGAIVYAKGLVILGNGGAEFGARG-FVSAFDAETGKLKMFYTVPNKKPEPDHAASDNL 256
 Db 228 ----EWMGDFEARMWT-----GYWGOITPDPVTMLVFSGVCGVPSASTGCTGGTL 277
 QY 257 MKKAYKTMGP--KSAWVRQGGGGTWDSDLVDPVSDLIYAVANGSPMNKYSSEGISN 314
 Db 278 YGNTREPAVPDTCGEIYVRHQTLPRDNWDOECTFEMKVVANVDQPSAWEGLRANPMA 337
 QY 315 LFLGSIYALKPEICEYVWHFQATPMDQDYSVQQTITLMPYK----- 358
 Db 338 TGE-RRLTGAPCKTGMSPDASGEFLMARDNTNMTASIDE-TGLVTVEDAVLKE 395
 QY 359 -GEMRHVYHAP-KNGFVYLDKATGFEISGXNYVQNMANGLDPLTGRPMYNDGLYTL 416
 Db 396 LDVEYDVCPTFLGGRDWSAALNPDGTIYFLPLNACVDI-----MAVDQ 440
 QY 417 NGKFWYIGPGLAHNMMAVYSPKTHLYIIPAHQIPFGYKNQVGFEPKPHADSNVGLDM 476
 Db 441 EFSALDYNTSATAKLAPEFNNGRIDALDITSGRTLSAERPAVNSPVLTAGVVEN 500
 QY 477 TKGNLPDTPPARAYIK---DLHGWLAMPVMEITVWKIDHKGPMNGGILATGDLTFQ 533
 Db 501 GGTDREYFALSOETGETLMQARLATVATGQAISEYLDGVQYIA--IGAGSL 549
 QY 534 GLANGEPHADATNGSDLYKFDAGSGIAPPTIYVNGQYVAVEVWGSI 584

RESULT 15
 ADD24941
 ID ADD24941 standard; protein; 803 AA.
 XX
 AC ADD24941;
 XX
 DT 15-JAN-2004 (first entry)
 XX
 DE Escherichia coli glucose dehydrogenase.
 XX
 KM Directed genetic engineering; galactomannanase; reduced activity;
 KM enhanced activity; xanthan gum production; suspension stability;
 KM emulsion stability; temperature resistance; pseudoplasticity; amylase;
 KM cellulase; extracellular protease; intracellular protease;
 KM glucose dehydrogenase; enzyme.
 XX
 OS Escherichia coli.
 XX
 PN US2003036176-A1.
 XX
 PD 20-FEB-2003.
 XX
 PF 10-AUG-2001; 2001US-00927827.
 XX
 PR 28-MAR-2001; 2001US-0279493P.
 XX
 PA (BOWE/) BOWER S G.
 PA (RAMS/) RAMSEIER T M.
 XX
 PI Bower SG, Ramseier TM;
 XX
 DR WPI: 2003-625389/59.
 DR N-PSDB; ADD24915.
 XX
 PT New transformed cell or organism having reduced or enhanced activity of
 at least one protein, useful for producing xanthan gum, which are useful

PT for providing formulations and properties.
 PS Claim 1; SEQ ID NO 69; 135pp; English.
 XX
 The present invention relates to polypeptide and polynucleotide sequences from Xanthomonas campestris which may be used for activity reduction or enhancement using directed genetic engineering. A transformed cell or organism having reduced or enhanced activity of at least one such protein e.g. galactomannanase can be generated by disrupting the gene encoding the protein. The activity of the protein is reduced by the presence of an antisense nucleic acid sequence. The nucleic acid sequence of the gene encoding the protein is a recombinant sequence having at least one mutation as compared to the wild-type gene encoding the protein. The transgenic cell or microorganism are useful for producing xanthan gum, which are useful for providing formulations and properties, such as long-term suspension and emulsion stability in alkaline, acid, and salt solutions, temperature resistance, and pseudoplasticity. The present sequence represents an enzyme relating to the present invention.

SO Sequence 803 AA;

Query Match 13.4%; Score 412; DB 7; Length 803;

Best Local Similarity 23.8%; Pred. No. 1.1e-24;

Matches 179; Conservative 67; Mismatches 239; Indels 266; Gaps 29;

QY 18 ABAFAQVTPITDELLANPPAG-----EWINVRNENRHSPLTQITADNV 64
 Db 129 AEGALTQVAGIAGAVNRPADQAPAPAPVDDDMAYGRSQGRISPLTQINRDNV 188
 QY 65 GOLQVWARGMEAGAV-----QVTPMINDGVMLANPBGDVIQALDAQDGLIWEH--- 114
 Db 189 SOLQQAMV--FHTGDVPFKRWGAETTLKVGDSLYLCSARNOVIALDASTQKQRRYDPK 246
 QY 115 --RQQLPAVILNAGDKRKRYALY-----GSLYFSSSDN 148
 Db 247 VKDEAIPYTAAC-----RGVSYEVEPTAAPAANDAAQAAPALCRFRVIEGTLDG 299
 QY 149 HLIALDVEY-----GQVPEVERSGEDGITSNTTGPVANGVIVAGSTC----- 193
 Db 300 RLIALDARTGPPCADFGTHGVDITVGMGERPPGVISNSPALYRGVVVIGHQVLDQK 359
 QY 194 QYSPYGCFTISGHDATGSELM-----RNHFTIPOGSEBDETMGNDEPAAMTGVWQO- 245
 Db 360 RYBPSG-VIOGFDAVTGLRWAMDTPENNGAPPG---QITWRTGPNMTTAADEQ 414
 QY 246 --IITDPTVNLV--FYGSTGVGPASSETQGTGGTLVGTNTRFVVRPDTGEIYVRHQTL 301
 Db 415 LGYVILPMGNSTADYSSS-----RTPENRRTLSL-VALDVTGKPVNRFQTH 463
 QY 302 RDNDQECTFEMVYANVDQPSAWEGLRANPNAATGEREVLTPAPCKTGMWSPDAS 361
 Db 464 IDAMDYD-----LGSQSLIDFPDQGVNPAVL-----LPSKQELVYLDERT 506
 QY 362 GEFLMARDNTNMTASIDETGLTVNEDAV-----LKE 395
 Db 507 GE-----PLVGEERAVPGGVEPQKRYKTOPPSLYHTLKKPDLTE 547
 QY 396 LD-----VEYDVC-----PTFLGARDWSAALNPDGTIY 424
 Db 548 RDMWGMPIDQLVCRIGPKASVYKIYTPPEADSHSIEYFPYNGSGDVGSAVDPHKGVI 607
 QY 425 FLPLAN-ACYDIM-----AVDQFSAIDVYNTSATAKLA 457
 Db 608 VANYNDMPNINILVPRAKADKLGNAPRDQVGDAGAGADPGATPYAVANVAGWRLP 667
 QY 458 PGFENN-----GRIDAIDISTGRTLSAERP-----AANYSP 489
 Db 668 --FTMLCKQPPYGGIRALIDKSGKITL--DRPFGSARGNPPFIRGSLPIETGPNNG 723
 QY 490 VLSTAGVVP--NGTDRYFALSOETGETLMQARLATVATGQAISEYLDGVQYIAGAG 548
 Db 724 SVTVAGGLIFFAAATVDDIRAIIDLATGKEIWHAKLPAGGQANPMVYVYAGGREYLVIMAG 783

Qy 549 LTYGTQJNAPLAEIDSTSVGNAIYVFPQ 579
Db 784 -----HHFMTPPAGDALVAVYALPQ 802

Search completed: November 18, 2004, 05:16:09
Job time : 165 secs

GenCCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: November 18, 2004, 05:13:08 ; Search time 40 Seconds
(without alignments)
959.953 Million cell updates/sec

Title: US-10-802-682-8
Perfect score: 3069
Sequence: 1 MNPFTLLRSTAAVLLTTPA.....AAIDSTSVGNALYVAPLQ 579

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCUTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/BACKFILE1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3069	100.0	579	US-09-470-667-8	Sequence 8, Appl1
2	2611	85.1	579	US-09-470-667-5	Sequence 5, Appl1
3	2526.5	82.3	579	US-09-470-667-6	Sequence 6, Appl1
4	2494	81.3	578	US-09-470-667-7	Sequence 7, Appl1
5	585	19.1	754	US-09-296-284-4	Sequence 4, Appl1
6	573.5	18.7	720	US-09-296-284-25	Sequence 25, Appl1
7	552	18.0	685	US-09-252-991A-32033	Sequence 32033, A
8	551.5	18.0	738	US-07-985-458-3	Sequence 3, Appl1
9	380.5	12.4	803	US-09-543-681A-4312	Sequence 4312, Ap
10	375.5	12.2	818	US-09-328-352-5208	Sequence 5208, Ap
11	367	12.0	740	US-09-635-145A-2	Sequence 2, Appl1
12	359.5	11.7	739	US-09-136-251-2	Sequence 2, Appl1
13	359.5	11.7	739	US-09-634-496-2	Sequence 2, Appl1
14	341.5	11.1	826	US-09-328-352-7515	Sequence 7515, Ap
15	302	9.8	602	US-09-489-039A-12436	Sequence 12436, A
16	286	9.3	688	US-09-489-039A-9813	Sequence 9813, Ap
17	216.5	7.1	706	US-09-252-991A-25730	Sequence 25730, A
18	161.5	5.3	243	US-09-252-991A-25814	Sequence 25814, A
19	154.5	5.0	339	US-09-543-681A-7743	Sequence 7743, Ap
20	133.5	4.3	335	US-09-540-236-2537	Sequence 2537, Ap
21	131	4.3	621	US-09-311-626B-2	Sequence 2, Appl1
22	126	4.1	715	US-09-710-279-570	Sequence 570, App
23	125	4.1	401	US-09-489-039A-11546	Sequence 11546, A
24	123	4.0	846	US-09-134-001C-3067	Sequence 3067, Ap
25	122	4.0	622	US-09-311-626B-4	Sequence 4, Appl1
26	122	4.0	1721	US-08-700-651-5	Sequence 5, Appl1
27	121.5	4.0	445	US-09-252-991A-22912	Sequence 22912, A

28	120.5	3.9	2736	4	US-09-252-991A-30227	Sequence 30227, A
29	120	3.9	910	4	US-09-710-279-3108	Sequence 3108, Ap
30	119.5	3.9	1042	3	US-08-928-361B-11	Sequence 11, Appl
31	119.5	3.9	1042	4	US-09-588-995A-11	Sequence 11, Appl
32	119.5	3.9	1837	3	US-08-928-361B-5	Sequence 5, Appl1
33	119.5	3.9	1837	4	US-09-588-995A-5	Sequence 5, Appl1
34	119	3.9	1043	3	US-08-928-361B-30	Sequence 30, Appl
35	119	3.9	1721	3	US-08-928-361B-6	Sequence 6, Appl1
36	119	3.9	1721	4	US-09-588-995A-6	Sequence 6, Appl1
37	118.5	3.9	470	4	US-09-710-279-2496	Sequence 2496, Ap
38	114	3.7	462	4	US-09-252-991A-29038	Sequence 29038, A
39	112	3.6	3567	2	US-07-642-734C-4	Sequence 4, Appl1
40	112	3.6	3567	3	US-08-439-009A-4	Sequence 4, Appl1
41	111	3.6	1183	3	US-09-134-001C-3530	Sequence 3530, Ap
42	110.5	3.6	383	3	US-09-134-001C-3068	Sequence 3068, Ap
43	110.5	3.6	1099	4	US-09-881-654-4	Sequence 4, Appl1
44	110.5	3.6	1099	4	US-10-637-323-4	Sequence 4, Appl1
45	110	3.6	3892	4	US-09-328-352-5503	Sequence 5503, Ap

ALIGNMENTS

RESULT 1	US-09-470-667-8	Application US/09470667
Sequence 8, Appl1	US-09-470-667-8	Patent No. 6730503
GENERAL INFORMATION:		
APPLICANT:	Asakura, Akira	
APPLICANT:	Hoshino, Tatsuo	
APPLICANT:	Ojima, Setsuko	
APPLICANT:	Shinjo, Masako	
APPLICANT:	Tomiyama, No. 6730503:bumi	
TITLE OF INVENTION:	No. 6730503:Alcohol/Aldehyde Dehydrogenases	
FILE REFERENCE:	C38435/109700CON	
CURRENT APPLICATION NUMBER:	US/09/470, 667	
CURRENT FILING DATE:	1999-12-22	
PRIOR APPLICATION NUMBER:	US 08/934,506	
PRIOR FILING DATE:	1997-09-19	
NUMBER OF SEQ ID NOS:	12	
SOFTWARE:	Patentin version 3.1	
SEQ ID NO. 8		
LENGTH:	579	
TYPE:	PRT	
ORGANISM:	Glucobacter oxydans	
FEATURES:		
NAME/KEY:	SIGNAL	
LOCATION:	(1)..(23)	
OTHER INFORMATION:		
US-09-470-667-8		
Query Match	100.0%; Score 3069; DB 4; Length 579;	
Best Local Similarity	100.0%; Pred. No. 2.8e-279;	
Matches	579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MNPFTLLRSTAAVLLTTPAFAQVTPITDELLANPPGEMINNGRQENRHSPLTQIT 60	
DB	1 MNPFTLLRSTAAVLLTTPAFAQVTPITDELLANPPGEMINNGRQENRHSPLTQIT 60	
QY	61 ADVNGLQVWARGMEAGAVOVTPVTHDGVWVLANPQGVICALDAQOTSDLIWHRRLCPA 120	
DB	61 ADVNGLQVWARGMEAGAVOVTPVTHDGVWVLANPQGVICALDAQOTSDLIWHRRLCPA 120	
QY	121 VATLNAQCDRRRGVALYGTSLYFSSGMDNHLTALDMETGQVVDVERSGEDGLTSNTTGP 180	
DB	121 VATLNAQCDRRRGVALYGTSLYFSSGMDNHLTALDMETGQVVDVERSGEDGLTSNTTGP 180	
QY	141 IVANGVIVAGSTCOVSPVGCFTSGHDSATGEBLNRNHIPOGEGDGTWGNDEPEARMT 240	
DB	141 IVANGVIVAGSTCOVSPVGCFTSGHDSATGEBLNRNHIPOGEGDGTWGNDEPEARMT 240	
QY	241 GWMGQITDPTNIVFYGSTVGPASFTORGTGPGTIVGNTFRVAPRDTEIYWRQTL 300	
DB	241 GWMGQITDPTNIVFYGSTVGPASFTORGTGPGTIVGNTFRVAPRDTEIYWRQTL 300	

Db 241 GVMGQITVDVPTNLVHVGSTGVGPASSTORGTGPGTLYGTNTRFAVRPDTGELVWEHQTL 300
Qy 301 PRDNWDECTFEEMVAVNDVQPSAEMEGRLAIPNNAATGERRVLTGAPCKGTGMSFDA 360
Db 301 PRDNWDECTFEEMVAVNDVQPSAEMEGRLAIPNNAATGERRVLTGAPCKGTGMSFDA 360
Qy 361 SGEFLMARDNTYNTMNASIDETGLVTNEDAVLKEIDVEDVCPPTFLGGRDMSAALND 420
Db 361 SGEFLMARDNTYNTMNASIDETGLVTNEDAVLKEIDVEDVCPPTFLGGRDMSAALND 420
Qy 421 TGIYFLPLNNACVDYMAVDEFSALDVYNTSATAKLAPFENNGRIDAIDISTGRTLMSA 480
Db 421 TGIYFLPLNNACVDYMAVDEFSALDVYNTSATAKLAPFENNGRIDAIDISTGRTLMSA 480
Qy 481 ERPAANSPVLTSTAGGVFENGTDTRYFRALSOETGETLWQARLATVATGQAISYEIDGQ 540
Db 481 ERPAANSPVLTSTAGGVFENGTDTRYFRALSOETGETLWQARLATVATGQAISYEIDGQ 540
Qy 541 YVIAIGAGLTYGTQNLNAPLA-EAIDSTSVGNATVVFALPQ 579
Db 541 YVIAIGAGLTYGTQNLNAPLA-EAIDSTSVGNATVVFALPQ 579

RESULT 2

US-09-470-667-5
; Sequence 5, Application US/09470667
; Patent No. 6730503
; GENERAL INFORMATION:
; APPLICANT: Asakura, Akira
; APPLICANT: Hoshino, Tatsuo
; APPLICANT: Ojima, Setsuko
; APPLICANT: Shinjoh, Masako
; APPLICANT: Tomiyama, No. 6730503ibumi
; TITLE OR INVENTION: No. 6730503el Alcohol/Aldehyde Dehydrogenases
; FILE REFERENCE: C38435/109700CON
; CURRENT APPLICATION NUMBER: US/09/470,667
; CURRENT FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: US 08/934,506
; PRIOR FILING DATE: 1997-09-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Gluconobacter oxydans
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(23)
; OTHER INFORMATION:
US-09-470-667-5

Query Match 85.1%; Score 2611; DB 4; Length 579;
Best Local Similarity 82.1%; Pred. No. 2,9e-236;
Matches 476; Conservative 51; Mismatches 51; Indels 2; Gaps 2;

Qy 1 MNPPTLIRTSAAVLLLTAPAPFAQVPTITDELLANPPAGMIVNGNOENYRHSPLTOIT 60
Db 1 MKPTSLIMASAGALALALAPAPFAQVPTITDELLANPPAGMIVNGNOENYRHSPLTOIT 60
Qy 61 ADVNGQLVWARGMEAGAVQVTPMTHDGVWYLANPGDVLOALDAQTGDLIMEHRQLPA 120
Db 61 TENVGQLVWARGMEAGAVQVTPMTHDGVWYLANPGDVLOALDAQTGDLIMEHRQLPN 120
Qy 121 VATLNAQGRKRKVALYGTSLYFSSWMDNHLIALDMETGOVFPVERSGSDGLTSNTGP 180
Db 121 VATLNAQGRKRKVALYGTSLYFSSWMDNHLIALDMETGOVFPVERSGSDGLTSNTGP 180
Qy 181 IVANGVIVAGSTCOYSPYGCIFSGHDSATGEELMRNHFIPQGEEDGTGNDFEARMT 240
Db 181 IVANGVIVAGSTCOYSPYGCIFSGHDSATGEELMRNHFIPQGEEDGTGNDFEARMT 240
Qy 241 GVMGQITVDVPTNLVHVGSTGVGPASSTORGTGPGTLYGTNTRFAVRPDTGELVWEHQTL 300
Db 241 GVMGQITVDVPTNLVHVGSTGVGPASSTORGTGPGTLYGTNTRFAVRPDTGELVWEHQTL 300

Db 240 GVMGQITVDVPTNLVHVGSTGVGPASSTORGTGPGTLYGTNTRFAVRPDTGELVWEHQTL 239
Qy 301 PRDNWDECTFEEMVAVNDVQPSAEMEGRLAIPNNAATGERRVLTGAPCKGTGMSFDA 360
Db 301 PRDNWDECTFEEMVAVNDVQPSAEMEGRLAIPNNAATGERRVLTGAPCKGTGMSFDA 360
Qy 361 SGEFLMARDNTYNTMNASIDETGLVTNEDAVLKEIDVEDVCPPTFLGGRDMSAALND 420
Db 361 SGEFLMARDNTYNTMNASIDETGLVTNEDAVLKEIDVEDVCPPTFLGGRDMSAALND 420
Qy 421 TGIYFLPLNNACVDYMAVDEFSALDVYNTSATAKLAPFENNGRIDAIDISTGRTLMSA 480
Db 421 TGIYFLPLNNACVDYMAVDEFSALDVYNTSATAKLAPFENNGRIDAIDISTGRTLMSA 480
Qy 481 ERPAANSPVLTSTAGGVFENGTDTRYFRALSOETGETLWQARLATVATGQAISYEIDGQ 540
Db 481 ERPAANSPVLTSTAGGVFENGTDTRYFRALSOETGETLWQARLATVATGQAISYEIDGQ 540
Qy 541 YVIAIGAGLTYGTQNLNAPLA-EAIDSTSVGNATVVFALPQ 579
Db 541 YVIAIGAGLTYGTQNLNAPLA-EAIDSTSVGNATVVFALPQ 579

RESULT 3

US-09-470-667-6
; Sequence 6, Application US/09470667
; Patent No. 6730503
; GENERAL INFORMATION:
; APPLICANT: Asakura, Akira
; APPLICANT: Hoshino, Tatsuo
; APPLICANT: Ojima, Setsuko
; APPLICANT: Shinjoh, Masako
; APPLICANT: Tomiyama, No. 6730503ibumi
; TITLE OR INVENTION: No. 6730503el Alcohol/Aldehyde Dehydrogenases
; FILE REFERENCE: C38435/109700CON
; CURRENT APPLICATION NUMBER: US/09/470,667
; CURRENT FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: US 08/934,506
; PRIOR FILING DATE: 1997-09-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Gluconobacter oxydans
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(23)
; OTHER INFORMATION:
US-09-470-667-6

Query Match 82.3%; Score 2526.5; DB 4; Length 579;
Best Local Similarity 79.8%; Pred. No. 2,5e-228;
Matches 462; Conservative 56; Mismatches 60; Indels 1; Gaps 1;

Qy 1 MNPPTLIRTSAAVLLLTAPAPFAQVPTITDELLANPPAGMIVNGNOENYRHSPLTOIT 60
Db 1 MKPTSLIMASAGALALALAPAPFAQVPTITDELLANPPAGMIVNGNOENYRHSPLTOIT 60
Qy 61 ADVNGQLVWARGMEAGAVQVTPMTHDGVWYLANPGDVLOALDAQTGDLIMEHRQLPA 120
Db 61 TENVGQLVWARGMEAGAVQVTPMTHDGVWYLANPGDVLOALDAQTGDLIMEHRQLPN 120
Qy 121 VATLNAQGRKRKVALYGTSLYFSSWMDNHLIALDMETGOVFPVERSGSDGLTSNTGP 180
Db 121 VATLNAQGRKRKVALYGTSLYFSSWMDNHLIALDMETGOVFPVERSGSDGLTSNTGP 180
Qy 181 IVANGVIVAGSTCOYSPYGCIFSGHDSATGEELMRNHFIPQGEEDGTGNDFEARMT 240
Db 181 IVANGVIVAGSTCOYSPYGCIFSGHDSATGEELMRNHFIPQGEEDGTGNDFEARMT 240
Qy 241 GVMGQITVDVPTNLVHVGSTGVGPASSTORGTGPGTLYGTNTRFAVRPDTGELVWEHQTL 300
Db 241 GVMGQITVDVPTNLVHVGSTGVGPASSTORGTGPGTLYGTNTRFAVRPDTGELVWEHQTL 300

Db 240 GWMGOITDYDVGGLVHYGSSAVGAPASETORGTGGTWTGNTFRFAVRPEIGETIWMHQTL 299
 QY 301 PRDWDDECTFEEMVAVNDVQPSAEMEGRLAIPNNAATGERRVLTGAPCKGTWMSFDA 360
 Db 300 PRDWDDECTFEEMVAVNDVQPSAEMEGRLAIPNNAATGERRVLTGAPCKGTWMSFDA 359
 QY 361 SGEFLMARDNTYNNMIASIDETGLTVNEDAVLKELDVEYDVCPTFLGGRDMSAALNDP 420
 Db 360 TGEFLMARDNTYNNMIASIDETGLTVNEDAVLKELDVEYDVCPTFLGGRDMSAALNDP 419
 QY 421 TGIYFLPLNACDYMADDOESFALDVYNTSATKLAPGFENMRIDAIDISTGRTLMSA 480
 Db 420 TGIYFLPLNACDYMADDOESFALDVYNTSATKLAPGFENMRIDAIDISTGRTLMSA 479
 QY 481 ERPAANSPLYSTAGVVFNGGTDYFRFALSQETGETLMQARLATVATGAQAISEYLDGVQ 540
 Db 480 ERPAANSPLYSTAGVVFNGGTDYFRFALSQETGETLMQARLATVATGAQAISEYLDGVQ 539
 QY 541 YVIAAGGLTYGTQNLAPLAERIDSTSVGNAIYVFPALPQ 579
 Db 540 YVIAAGGLTYGTQNLAPLAERIDSTSVGNAIYVFPALPQ 578

RESULT 4
 US-09-470-667-7
 ; Sequence 7, Application US/09470667
 ; Patent No. 6730503
 ; GENERAL INFORMATION:
 ; APPLICANT: Asakura, Akira
 ; APPLICANT: Hoshino, Tatsuo
 ; APPLICANT: Ojima, Setsubo
 ; APPLICANT: Shinjoh, Masako
 ; APPLICANT: Tomiyama, No. 6730503Jbun1
 ; TITLE OF INVENTION: No. 6730503el Alcohol/Aldehyde Dehydrogenases
 ; FILE REFERENCE: C38435/109700CON
 ; CURRENT APPLICATION NUMBER: US/09/470,667
 ; PRIORITY FILING DATE: 1999-12-22
 ; PRIORITY APPLICATION NUMBER: US 08/934,506
 ; PRIORITY FILING DATE: 1997-09-19
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: Patent In version 3.1
 ; SEQ ID NO 7
 ; LENGTH: 578
 ; TYPE: PRT
 ; ORGANISM: Gluconobacter oxydans
 ; FEATURE:
 ; NAME/KEY: SIGNAL
 ; LOCATION: (1)..(23)
 ; OTHER INFORMATION:
 ; US-09-470-667-7

Query Match 81.3%; Score 2494; DB 4; Length 578;
 Best Local Similarity 79.3%; Pred. No. 2.9e-225; Indels 2; Gaps 2;
 Matches 459; Conservative 56; Mismatches 62;
 QY 1 MNPFTLLRTSAVALLTAPAFQVTPITDELLANPPAGEMINYGNOENYRHSPLTQIT 60
 Db 1 MKLFTLLQSSAALLVLGTIPALAQ-TAITDEMLANPPAGEMINYGNOENYRHSPLTQIT 59
 QY 61 ADVNGOQLVYARMEGAVQVTPMIDGMYLANPBDVIOALDAQGDLIMEHRQLA 120
 Db 60 ADVNGOQLVYARMEGAVQVTPMIDGMYLANPBDVIOALDAQGDLIMEHRQLA 119
 QY 121 VATLNAQGRKRGVALYGTSLYFSSWNHILALDMETGQVVFVERSGEDGLTSYTTGP 180
 Db 120 VATLNAQGRKRGVALYGTSLYFSSWNHILALDMETGQVVFVERSGEDGLTSYTTGP 178
 QY 181 IVANGVYAGSTCOYSPGCTISGHSATGSELMRNHFIPOGHEGSETWGNPPEARWMT 240
 Db 179 IVANGVYAGSTCOYSPGCTISGHSATGSELMRNHFIPOGHEGSETWGNPPEARWMT 238
 QY 241 GWMGOITDYDVTNLYVFGSTGVGPASETORGTGGTWTGNTFRFAVRPEIGETIWMHQTL 300
 Db 240 GWMGOITDYDVGGLVHYGSSAVGAPASETORGTGGTWTGNTFRFAVRPEIGETIWMHQTL 299

Db 239 GWMGOITDYDVGGLVHYGSSAVGAPASETORGTGGTWTGNTFRFAVRPEIGETIWMHQTL 298
 QY 301 PRDWDDECTFEEMVAVNDVQPSAEMEGRLAIPNNAATGERRVLTGAPCKGTWMSFDA 360
 Db 299 PRDWDDECTFEEMVAVNDVQPSAEMEGRLAIPNNAATGERRVLTGAPCKGTWMSFDA 358
 QY 361 SGEFLMARDNTYNNMIASIDETGLTVNEDAVLKELDVEYDVCPTFLGGRDMSAALNDP 420
 Db 360 TGEFLMARDNTYNNMIASIDETGLTVNEDAVLKELDVEYDVCPTFLGGRDMSAALNDP 418
 QY 421 TGIYFLPLNACDYMADDOESFALDVYNTSATKLAPGFENMRIDAIDISTGRTLMSA 480
 Db 420 TGIYFLPLNACDYMADDOESFALDVYNTSATKLAPGFENMRIDAIDISTGRTLMSA 478
 QY 481 ERPAANSPLYSTAGVVFNGGTDYFRFALSQETGETLMQARLATVATGAQAISEYLDGVQ 540
 Db 480 ERPAANSPLYSTAGVVFNGGTDYFRFALSQETGETLMQARLATVATGAQAISEYLDGVQ 538
 QY 541 YVIAAGGLTYGTQNLAPLAERIDSTSVGNAIYVFPALPQ 579
 Db 540 YVIAAGGLTYGTQNLAPLAERIDSTSVGNAIYVFPALPQ 577

RESULT 5
 US-09-296-284-4
 ; Sequence 4, Application US/09296284A
 ; Patent No. 6204040
 ; GENERAL INFORMATION:
 ; APPLICANT: Choi, Bui-Sung
 ; APPLICANT: Rhee, Sang-Ki
 ; APPLICANT: Lee, Eun-Hae
 ; TITLE OF INVENTION: Gluconobacter Suboxydans Sorbitol Dehydrogenase, Genes
 ; FILE REFERENCE: 1533.0870000
 ; CURRENT APPLICATION NUMBER: US/09/296,284A
 ; PRIORITY FILING DATE: 1999-04-22
 ; NUMBER OF SEQ ID NOS: 87
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 4
 ; LENGTH: 754
 ; TYPE: PRT
 ; ORGANISM: Gluconobacter suboxydans
 ; US-09-296-284-4

Query Match 19.1%; Score 585; DB 3; Length 754;
 Best Local Similarity 28.2%; Pred. No. 9e-46; Indels 106; Gaps 21;
 Matches 172; Conservative 94; Mismatches 238;
 QY 6 LRTSAV-LLTAPAFQ-VTPITDELLANPPAGEMINYGNOENYRHSPLTQITAD 62
 Db 16 LIGCAALAFCATSPVALAEDTGTATN-ADQHPGDMMSGRTYSBQRXSPLDQITKD 72
 QY 63 NVGOQLVW-----ARGMEGAVQVTPMIDGMYLANPBDVIOALDAQGDLIMEHR 116
 Db 73 NASNKLAMHYDLITNRQEG-----TPLIVGVVATTNMSKMLALAAQKLMSTDP 127
 QY 117 QLPV-VATLNAQGRKRGVALYGTSLYFSSWNHILALDMETGQVVFV----- 164
 Db 128 KVPGNIAIRGCCDVTNRGAAYWNGKVYGTGRLIALDATGKLVSVYVTPKXAGLGH 187
 QY 165 ERSGEDGLTSYTTGPYANG-VIYAGSTCOYSPGCTISGHSATGSELMRNHFIPOG 223
 Db 188 QRSYTVDG-----APRIAKGVIIINGGAFAAG-FTVAYDDETGMDDRRFTVPDP 240
 QY 224 EGGD-----ETWGNDFEARWMTG-----VWGQITDYDVTNLYVFGSTGVGPASE 267
 Db 241 NKPDAAADDVLMKATYPTWKGAMKQGGGCTVMDLIDPVTDLVYLGVNGSPNNY 300
 QY 268 TORGTGGTLYGTNTRFAVRPEIGETIWMHQTLPRDWDDECTFEEMVAVNDVQPSAEME 327
 Db 301 KFRSGKGNLFLGIVAINPDTGKYVWFQETPMDDMDYTSVOQIVA--LDMPVNGEM- 357
 QY 328 GLRAINPNNAATGERRVLTGAPCKGTWMSFDAAGEFLMARDNTYNNMIASIDE-TGLVT 386

Db 358 -----RHVYHAP-XNGFFYIIDATGKFRISGRPTYENMANGLDPVYGRN 403
QY UNEDAVLKEHDEVEDVPTFLGGRDWSAALNPDQIYFLPLNNACYDIAMAVDOESALD 446
Db 404 YNBDALMTLNGKRWYGIPLGLGHNFAAMVAPQTKLVYIPAOQVF---VYDPQKGRK 460
QY 447 VYNTSATKLAPEGENNRIIDAIDISTGR-----TLMSAREPAAVNS-- 488
Db 461 AHHDSWNLGL---DNKKGILLDDNDPOHKADKQFLDKGMIVAMPQXQQAFTVD 515
QY 489 -----PVLTAGGVFENGSTDRYFALASQETGETLMOARLATVATGQAISEYLDGVQY 541
Db 516 HKBPMMGLATAGGVILFQGLANGEFHAYDATTKDLFTFPASALIAAPVITYANGKQY 575
QY 542 IA--IGAGGL 549
Db 576 VAVEVGWGI 585

RESULT 6
US-09-296-284-25
Sequence 25, Application US/09296284A
Patent No. 6204040
GENERAL INFORMATION:
APPLICANT: Choi, Eun-Sung
APPLICANT: Rhee, Sang-Ki
APPLICANT: Lee, Eun-Hae
TITLE OF INVENTION: Glucobacter Suboxydans Sorbitol Dehydrogenase, Genes
FILE REFERENCE: 1533.0870000
CURRENT APPLICATION NUMBER: US/09/296,284A
CURRENT FILING DATE: 1999-04-22
NUMBER OF SEQ ID NOS: 87
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 25
LENGTH: 720
TYPE: PRT
ORGANISM: Glucobacter suboxydans
US-09-296-284-25

Query Match 18.7%; Score 573.5; DB 3; Length 720;
Best Local Similarity 27.9%; Pred. No. 1e-44;
Matches 164; Conservative 90; Mismatches 230; Indels 103; Gaps 19;
QY 26 TPITDELLANPAGEMINYNQENYRHSPLTQITADNVGOLQVW-----ARGWEAGA 79
Db 5 TAITN---ADQHGDMMSYGRVTSSEGRYSPLDQITDMNSNLKLAHMYLDITRQGE-- 59
QY 80 VQVTPMHDGVTLANPBGVITQALDAQOTGLIWEHRRLPA-VATLNAQGRKGVALY 138
Db 60 ---TPLLVDGVYATTNWSKKKALDAATGKILMSYDPKVPGNADRGCCDTPVRGAAYN 116
QY 139 TSLYFSSWDHLIALDETGOVVDV-----ERGSGEDGLTNTGPIVANG-V 186
Db 117 GKTYFGFDRLIALDKTKGLVMSYITVPEKQLOHQSYYVDG-----APRIKGVY 170
QY 187 IVASGTQVSPYCGFISGHDATGEELRNHFIPOGEED-----ETWAND 233
Db 171 IIGGGALEFGARG-FVTAAYDAETGKMDWRFPTVPNDKPDGGAASDVILMSKAYPTWKG 229
QY 234 FEARMTG---VNGQIYDPVTNLVFGSTGVGPASETQKGTGGLYGTNTRFAVRPT 290
Db 230 GAWKQGGGGTVDLSLYDPVTLVYLGNGSPMYKFRSBSGKNNLFGSLVALNPT 289
QY 291 GEIWRHQLPRDNWDECTFEMNVAVNDVQPSAEKGLRATNPNAATGERVLTGAPCK 350
Db 290 GKTYVHFQETPMQWDYTSVQQLMA-LDWPVNGEM-----RHVLYHAP-X 332
QY 351 TGTMSDASGEFLMARDTNTNMKASIDB-TGLVTVNEDAVLKEHDEYDVCPFTLGG 409
Db 333 NGFFYIIDATGKFRISGRPTYENMANGLDPVYGRNPDALMTLNGKRWYGIPEDLGG 392

QY 410 RDMSSALNPDQIYFLPLNNACYDIAMAVDOESALVNTSATKLAPEGENNRIIDA 469
Db 393 HNFAMAVSPQTKLVYIPAOQVF---VYDPQKGGFRAHDSNMLG---DNKKGILL 444
QY 470 DISTGR-----TLMSAREPAAVNS-----PVLTAGGVFENGST 504
Db 445 DDNDPOHKADKQFLDKGMIVAMPQXQQAFTVDHKBPMMGLATAGGVILFQGLAN 504
QY 505 RYFPLASQETGETLMOARLATVATGQAISEYLDGVQYIA--IGAGGL 549
Db 505 GEFFHAYDATTKDLFTFPASALIAAPVITYANGKQYVAVEVGWGI 551

RESULT 7
US-09-252-991A-32033
Sequence 32033, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 32033
LENGTH: 685
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32033

Query Match 18.0%; Score 552; DB 4; Length 685;
Best Local Similarity 27.2%; Pred. No. 9.6e-43;
Matches 180; Conservative 110; Mismatches 230; Indels 142; Gaps 28;
QY 3 PTLTARTSA-----AVLLTPAFAQVPTITDELLAN--PPAGEMINYNQENYRHS 55
Db 71 PAGLRPSHLCLARAVALGSAALAK--DYTWEDINDKTKTGVDLYQWGTGAQRWSP 128
QY 56 LTQITADNVGOLQVWMA-----RQMEGAYOVYTMINDGVYTLANPBGVITQALDAQ 107
Db 129 LQVNVADNVFLTPAMWSYSPDEKQROESQAI-----VBDGYTVYASYSRLFALDAKT 183
QY 108 GDLIWEHRRLP-----PAVATLNAQGRKGVALYGTSLYFSSWDHLIALDETGOVVF 162
Db 184 GKRLMTYHRLPDIRCCDVN-----RQALYDGKVFPGTLDASVALNKNTGKVVW 237
QY 163 DVERGSGEDGLTNTGPIVANG---VIVAGST-CQYSPYCGFISGHDATGEELRN 216
Db 238 KKKFADHAGAYTMGAPTIYADGKTGKVLHIGSSGDFGVGLFA-RPDPGEEELMR 296
QY 217 HFIPOGEED---DETWANDFEA-----RMWTG---VNGQIYDPVTN 253
Db 297 PFV---EGHMERLNGKSDITVGDVAPSPWPPDRNSPTKVESMSHGGAAPQASFDLETN 354
QY 254 LVFGSTGVGPASETQKGTG---TLYGTNTRFAVRPTGEIWRHQLPRDNWQ 307
Db 355 TIIVAGANPGRPMNTWARTAGGNPHDYDSLX-TSGQGVDPSSGVEVCFYQHTPNDAMDF 413
QY 308 ECTEEMVAVNDVQPSAEMEGRLAINPNAATGERVLTGAPCKTGTMTSDDASGE--- 363
Db 414 SGNNELVY-----FDYKADGKIVKATAHADRNGFFVYVDRSGKQLQNA 457
QY 364 -----ELMARDTNTNMKASID-ETGLVTVNED--AVLKEHDEY---DYCPFTLGR 410
Db 458 FPFVDNTWMA-----SHIDLKGRPYERREGQRPPLPEQKQKGAVEGSPPLGK 508
QY 411 DWSSALNPDQIYFLPLNNACYDIAMAVDOESALDVNTSATKLAPEG-----ENM 463

; Sequence 4312, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709, 1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4312
; LENGTH: 803
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4312

Query Match 12.4%; Score 380.5; DB 4; Length 803;
Best Local Similarity 24.5%; Pred. No. 1,6e-26;
Matches 182; Conservative 87; Mismatches 240; Indels 233; Gaps 40;
QY 5 TLRTSAVLL--LTPAA---FAQVPTIDELLANPPAG---EWINGENQENYRH 53
DB 124 TLATTLIIIGSLTDPHAINGFQOCP-TNSI--SPGVNDNDWPAYGRTQAGIRY 179
QY 54 SPLTQITADNVGLOLVMA-----RGMEAGAV--QVTPMTHDGVMTLANPGDVIQALD 104
DB 180 SPLQIINSNDVNDKLVKMTPTHTGDEKANDAVEITHEVTIKIEDNIFLCTPHQYIISD 239
QY 105 AQTGDLIMEHRRQLPAVALINAQDREKRGVALYGTG-----L 141
DB 240 PATGEKMRPDSKIQ--YNSKSFQHMTCRGVAAYYNNNSAEFAKSLDVRQPSIDCEPEKY 297
QY 142 YFSSMDHLIALDMETQGVFDERGSGEDGLTS-----NTTGP--IYANVIYA 189
DB 298 FLVNDRLIAVNAQTNACRDRGQ--NGEINILASNPYALGGYNTPTSPYVIGNTIITIG 356
QY 190 GSTQC---YSPYCGFISGHSATGEELMNHFIPOGEE-----GDETGNDFEARWM 239
DB 357 GSVTDNLSGTQEPSC-VIRGYDVNSGKLIW---VPTGADPNAMGKET--QF-VHNS 407
QY 240 TGWGGITTDPRVTLNVYSGTGVGPASGTQGTG---GGTLVGTNTRF-----AVRPDGG 291
DB 408 PNANAPLAYDAQTDIV-YIPTGV-----GTPDIWGDGRHPLKERYANSVLALGSGTG 458
QY 292 EIVVRHQTLPDNDMDCEFEEMVAVNDVQPSA-----EMEGIRAINPMAATGERRVL- 344
DB 459 ELIMHFTQTHHDLMDND-----VPSQPSLVDIKTDKNETIPALYIYTKTGAVFVLD 509
QY 345 -----TGAPCKTGTMNS-----FDPAAS-----GEFLMA----- 367
DB 510 RRDGTPIVPYTERKVPSTVYKXGQTOGEHYSPTOPFSALNLAIRPNKLOSGDWMGGMADQ 569
QY 368 -----RDNYTMMIASIDETGLVTVNEDAVLKELDVEYDVCPFTLGGDRWSAALNDP 420
DB 570 LACRIAHQNLBSGITYPPSENGTL-----VFPGLUGIFEMWGISVND 613
QY 421 TGT-----YFLPLNNACY-----DIMAADQ--EFSALDVYNTSATRAKLAPEGEMNGR- 465
DB 614 RQVAVNAFITLPFISKYLLKPDQDNQALDAGHGBOGLQPMGTGTPYSVDIHFPISPLGLPC 673
QY 466 -----IDMIDISTGTLMS-----AERP-----AANYSPVLSAGGVN 500
DB 674 KQRPWGFVAGVDLTKNELVWQORFGTTRDSLEPNLVEPLKIGVPGIGGAISTAGVNVRY 733
QY 501 GGT-DRYFRALSOFTGETLMOARLATVATGQA--ISYELDGVYIAGAGL-TYGTQAIN 556
DB 734 AASGDNYIRAFNVNTNGMLWEARLP--AGGQATPMTYISGQYVIVMAGHGSGFGTKM- 790
QY 557 APLAEALDSTSVGNALIVPALP 578
DB 791 -----GSLVAYGLP 800

RESULT 10
US-09-328-352-5208
; Sequence 5208, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5208
; LENGTH: 818
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5208

Query Match 12.2%; Score 375.5; DB 4; Length 818;
Best Local Similarity 24.1%; Pred. No. 4,9e-26;
Matches 177; Conservative 87; Mismatches 253; Indels 217; Gaps 38;
QY 6 LLRTSAVLLTAPAAFAQVPTIDELLANPP-----AGEWINGENQENYHSP 55
DB 138 LSSTLAIIVLMVYSIFNDPQEIINGEIKTPQETAGAVPGVAESDWPAYGRTQAGERYSP 197
QY 56 LTOITADNVGLOLVMA-----RGMEAGAV--QVTPMTHDGVMTLANPGDVIQALDQ 106
DB 198 LKQINDQVNDKLVKMTPTHTGDEKANDAVEITHEVTIKIEDNIFLCTPHQYIISD 257
QY 107 TGDILMEHRRQLPAVALINAQDREKRGVALY-----GTSI-----YFS 144
DB 258 TGEKMRPDSKIQ--YNSKSFQHMTCRGVAAYYNNNSAEFAKSLDVRQPSIDCEPEKY 315
QY 145 SWDNLIALDMETQGVFDERGSGEDGLTS-----NTTGPVANG--YIVAGS- 191
DB 316 VNDGLVAVNADTGACTDFGQ--NGEVNLCFEMFYALPGYNTPTSPYVIGNTIITIG 374
QY 192 TCQYS---PYCGFISGHSATGEELM-----RNHFIPOGEGDETGNDFEARWM 240
DB 375 TDNYSNKEPSC-VIRGYDVNTGKLIWFDGADPN--AMPGB-----GTFP-VHNSP 423
QY 241 GVGGGITTDPRVTLNVYSGTGVGPASGTQGTG---GGTLVGTNTRF-----AVRPDGG 290
DB 424 NAMAPLAYDAQTDIV-YIPTGV-----GTP--DIWGDRTLEKERYANSVLALGSGTG 458
QY 291 GEIVVRHQTLPDNDMDCEFEEMVAVNDVQPSAEMEGIRAINP--NAATGERR----- 342
DB 473 GKLIWFQTHHDLMDNDVPSQPSLADIKKSGKTVAIYVLTGTGNAFVLDNRNGQPIV 532
QY 343 -----VLTGAPCKTGTMNS-----FDPAASGEFLMA-----R 368
DB 533 PTEREPVQTVKRGQOTKGEHYSTKTOPFSALNLAIRPNKLOSGDWMGGMADQ 569
QY 369 DTNYTMMIASIDETGLVTVNEDAVLKELDVEYDVCPFTLGGDRWSAALNDP 420
DB 593 RLNDGQITTPSENGTL-----VFPGLUGIFEMWGISVND 613
QY 421 TGIVP---LPLN-NACYDIADVQEPSALDV---YNTSATYADP-GF---ENMGRI 466
DB 637 IGLPVSRLIPADPRAGTAKAGTEQGVQPMYGVVPGVEISAFSLSLGPKCPAWGVY 696
QY 467 DAIDISTGRTLSA-----ERPA-----NYSPTVLSAGGVNNGGT-DRYF 507
DB 697 AGVDLKTBEVWVKRIGITRDSLEPNLRLQPLPVKIGVGLGSGISTAGVNVFVATQNYI 756
QY 508 PALSQETGETLMOARLATVATGQA--ISYELDGVYIAGAGL-TYGTQAIN 556
DB 757 RAFNVNBEKLMWARLP--AGGQATPMTYISGQYVIVMAGHGSGFGTKM- 805
QY 565 STSVGNALIVPALP 578

QY 346 -----GAPC-----KGTGMSFPAASGEF--LMARDNTYNNMIA--SIDETGLVTVNEDA 391
DB 465 PMSVGMALRVDPDKETDMMGMSPIDOLFCKRIKFRANVGEFTPPSVDKPM----- 516
QY 392 VLKELDEYDVCPFLGGRDMSAALNPDGTGYELPLN--NACVD-----IMAV 438
DB 517 -----IEY-----PGYNGSDMSMSYDPQSGILIANMNTIPMTDQVTRKKADSLGLMPT 567
QY 439 D-----OEFSAIDVYNTSATKAPGEN-----MGRDAIDISTG-RT 476
DB 568 DDNFKKGGGAGENGAMD--GTPYGIIVTPFMDQYTKMCMNPPYGMITAIIDMKHQKV 625
QY 477 LMS-----AERP-----AANYSPVLSTAGGVENG-GTDRYFRALSOETGE 516
DB 626 LMQHPGLGTARANGPWGLPTGLPWEIGTPNNGSSVVTGGGLIFIGATIDNOIRALDEHTGX 685
QY 517 TLMQARLATVATGQATSYELDVOYIAGAGLTYGTQNLAPLAELIDSTVGNALVYFA 576
DB 686 VVMSAVLPFGGQANPMTYEANGHQYVAI-AGG-----HHFMPTVSDQLVYVA 732
QY 577 LP 578
DB 733 LP 734

RESULT 13

US-09-634-496-2
Sequence 2, Application US/09634496
Patent No. 644449
GENERAL INFORMATION:
APPLICANT: HOSHINO, Tatsuo
APPLICANT: MIYAZAKI, Taro
APPLICANT: OJIMA, Setsuko
APPLICANT: SHINJOH, Masako
APPLICANT: TOKIYAMA, No. 644449ibumi
TITLE OF INVENTION: D-Sorbitol Dehydrogenase Gene
FILE REFERENCE: D-Sorbitol Dehydrogenase Gene
CURRENT APPLICATION NUMBER: US/09/634,496
EARLIER FILING DATE: 2000-08-08
EARLIER APPLICATION NUMBER: US/09/136,251
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 739
TYPE: PRT
ORGANISM: Gluconobacter suboxydans
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)..(24)
US-09-634-496-2

Query Match 11.7%; Score 359.5; DB 4; Length 739;
Best Local Similarity 23.7%; Pred. No. 1.3e-24;
Matches 172; Conservative 83; Mismatches 229; Indels 239; Gaps 36;
QY 23 AQTPTIDELLANPBAEMINNGNQNENRHSPLQITLADNVOQLQ--VMAG----- 74
DB 86 SQVPAMAPQOSAPPARQDWAAYGDDHQRTYSPSETTEPENASKLKVAFYHTGTYPRQG 145
QY 75 -MEAGVQVTPM-LHDSVMTLANGVYIQAIDAQTDLIWEHR-----ROLPAVATLNAQ 127
DB 146 QVNMMALETPIKVGDDLYTCSAMNDIK-LDPATGQIWRRAVDVYKHSIFPTAAC--- 201
QY 128 GDRKGVALYGT-----LYSSWDNHLIADMETGVV-----FVDERGS 168
DB 202 ---KGTTFPTSSVPEGQPCENRLLEGTLDMRLIADVADTGFCEFGHGQVNLMOGL 257
QY 169 GED--GTSNTTGPVANGVYVAGSTC-----QYSPYGCFIGSDSATGEELW---RNH 217
DB 258 GESVPSVSWTAPRPVINGVAVVNHVLDQGRMAPEG-VIRGVDAESGKFWAMVNVNS 316

QY 218 FIPQGEEDWGNDEEARMWMTGVMOQITYPVNLVFGSTGVGPASETORGTGCTL 277
DB 317 GRGQPA-----YRVTVITAVERRIP-----GLPDRRGG----- 345
QY 278 YGINTR-----FAVRPD-----TGEIYWRHQTLPRDNDQDECTFE 312
DB 346 -GSRLEPRDRNSAADYSALRSADENKVSAAVAIDVKTSPRVFQTAHXDVNDYDGSQ 404
QY 313 ---MVAIVNDVQ-----PEAEMEGLRALINPAATGERRYLT- 345
DB 405 ATLMDMGPPDQIYFALIMPYKGGQTFVLDRTGKPLPVEERPAESPQVTPDDPSPTQ 464
QY 346 ---GAPC-----KGTGMSFPAASGEF--LMARDNTYNNMIA--SIDETGLVTVNEDA 391
DB 465 PMSVGMALRVDPDKETDMMGMSPIDOLFCKRIKFRANVGEFTPPSVDKPM----- 516
QY 392 VLKELDEYDVCPFLGGRDMSAALNPDGTGYELPLN--NACVD-----IMAV 438
DB 517 -----IEY-----PGYNGSDMSMSYDPQSGILIANMNTIPMTDQVTRKKADSLGLMPT 567
QY 439 D-----OEFSAIDVYNTSATKAPGEN-----MGRDAIDISTG-RT 476
DB 568 DDNFKKGGGAGENGAMD--GTPYGIIVTPFMDQYTKMCMNPPYGMITAIIDMKHQKV 625
QY 477 LMS-----AERP-----AANYSPVLSTAGGVENG-GTDRYFRALSOETGE 516
DB 626 LMQHPGLGTARANGPWGLPTGLPWEIGTPNNGSSVVTGGGLIFIGATIDNOIRALDEHTGX 685
QY 517 TLMQARLATVATGQATSYELDVOYIAGAGLTYGTQNLAPLAELIDSTVGNALVYFA 576
DB 686 VVMSAVLPFGGQANPMTYEANGHQYVAI-AGG-----HHFMPTVSDQLVYVA 732
QY 577 LP 578
DB 733 LP 734

RESULT 14

US-09-328-352-7515
Sequence 7515, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-039A
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 7515
LENGTH: 826
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-7515

Query Match 11.1%; Score 341.5; DB 4; Length 826;
Best Local Similarity 23.4%; Pred. No. 7.8e-23;
Matches 157; Conservative 83; Mismatches 215; Indels 217; Gaps 28;
QY 41 WINGRQENYRHSPLQITADNVGOLOLV-----AGMEAGAV-QYTPMIDGVMY 92
DB 190 WENQNDGSSRFVALDQINDNVHKLKEMRFRPTGFTTGSNGAGADQWTFLOVGKVL 249
QY 93 LANPQDVQALDAQTDLIWEHRROLPAVATLNAQD--RKRGVALYGT----- 140
DB 250 LCPHNNI.FAIDADSGKQIWK-----AEVNSKADAMERCAGVAYFDTKPLVQPTLA 301
QY 141 -----LYSSWDNHLIADMETGVV-----VERSGEGGLTSN- 176
DB 302 GATPVTVAANTACPRRYTNTPGRLIIVANVAGQSCADPGVNGVYDLLEGAGGSKAP 361
QY 177 ---TTGPIVANGVYVAGS-----TCQYSPYGCFIGSDSATGEELW---RN--HFIQ 221

```

Db      362 REEVTAFTAGTTIVGSRIDNVAADMFGVIRGVYITGLRMAFDPRNDPNVYK 421
Qy      222 PGESEDETGNDFEARMWMTGVMQITVDPVNTLVF--YSGTGVPASETGRTGPTLYG 279
Db      422 PGEIYKSSANS-----MAMSYDPQNTVTFLEPMSSSVYDINQGNR--NPLDKYN 470
Qy      280 TMTFAVAPDTGELVTHQTLPRDNMDOECTFEKCVANVDVQSAEKEGLRANPNAATG 339
Db      471 TSV-LAADATTGKEMVYQTVHNDLMD-----FDLPMQSLVDVFLK-----DGT 515
Qy      340 ERRLTGAPCKTGTMTGSEFDAAGEFLWABDTNVTNMI-----AS 378
Db      516 KPAVVIQT--KSGQFFVLDRVTGKPL-----TKVLEQPKAANTGEOYSLTOPRVE 566
Qy      379 IDETGLVTNEDAVLKELDVEYDVC-----PTFLGGRDMS 414
Db      567 MPOIGNOTLTRESDMGATPFDLMCRINFKSMRYEGFLTAGTDVSLSPFSGISGMMWS 626
Qy      415 AALNP-----DGIYFLPANNACYDIMAY 438
Db      627 IAPPTTRVYFVNDMLGLVQLIKQTPEDIKIQASGGEKVNTGKAVPMKGPYKVNK- 685
Qy      439 DOERSALDVYNTSATAKLAPFENMGRIDAIDISTGRTLMSAERPAAN-----Y 487
Db      686 NREMSAL-----SIPCKKP-----FGMTALDKMTKRVANOVPLGITVEDTGLIGKGLK 736
Qy      488 SPV-----LSTAGGVENGCT-DRYFPALSOETGETLMQARLATVATGAISTY--EL 536
Db      737 APIGMPTIGPMATOGGLVFPFAATQDYLLAFDSSTGKELKSMIPVSGGTISVSPK 796
Qy      537 DGVOYIAGAG 548
Db      797 TGKQYVVVTAG 808

RESULT 15
US-09-489-039A-12436
; Sequence 12436, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709,2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12436
; LENGTH: 602
; TYPE: PRY
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12436

```

Query Match 9.8%; Score 302; DB 4; Length 602;
 Best Local Similarity 23.3%; Pred. No. 2,4e-19;
 Matches 147; Conservative 81; Mismatches 214; Indels 190; Gaps 30;

```

Qy      81 QVPMHIDGVMYLANPBDVQALDAQTDGLIHEHROLPAVATLNAQDRKRGVALYGTIS 140
Db      23 EVTPIKVGDTLFLCTAHQRLFALDAATGKEMHFDPLQNDPSF-----QHVTGRVYS 75
Qy      141 LYFSSWDN-----HLIALDMETGVVFPVERGSGR----- 170
Db      76 YHEKADNAPADVADDCRRILIPVNDGR-LFAVNADNGKLCETPANKGILNQTMPVT 134
Qy      171 -DGLTSNTTGPYAN-GVIYAGS-TCQYS---PYGCFISGHSATGEBELRNHFTIQ--- 221
Db      135 TPQKYEPSTSPITTDKTIIVAGVTDNFSSTREPSG--VIRGFVNTGKTLMA--FDPGARD 191
Qy      222 -----PGESEDETGNDFEARMWMTGVMQITVDPVNTLVFGSTGVPASETGRTGPTL 277

```

```

Db      192 PNAISDEHHTLINS-----PNSMAPAAYDAKLDLVYL-----PMGVTPPIDWGNR 238
Qy      278 YGNTRF-----AVRPDGEIYVRHQTLPDWMDOECTFEKCVANVDVQSAEKEGLRAI 332
Db      239 TPEQERYASIVAAVATTKLAMSIVQTVHNDLMDNDMSQPTLADIEV----- 286
Qy      333 NPNMATGERRVLTGAPCKTGTMTSFPDASGEFL-----WARDTNY--TNMIASI 379
Db      287 -----NGKTVPVIVAPAKTGNIFVLDNRNGELVPAPEKVPQGAAGDYVTKTOPFSDL 341
Qy      380 -----DETGL-----VTNVEDAVLKE--DVEYD-----VCPPTFLGGRDMSA 415
Db      342 SFRPKDGLTADMGATMFDQVCRVITHQLEKYEIGITPPSEOGTLVFPNGIMFEMWGI 401
Qy      416 ALNPDTGIYF-----LPANNACY-----DINAVDOESALDVYNTSATAKLAP 458
Db      402 SYDPRQVAAIANPMLPFSKLIIRGPNPMPEPPKDAKSGTSGSVQPOYGVPIYVTLNP 461
Qy      459 GFEEM-----GRIDAIDISTGRTLMS-----AERPAANYSPVL--- 491
Db      462 FLSPFGLPCKOPAWCYISALDKTNEVYWKRICTPQDSLFPMPVKLPFTGKEMTLGP 521
Qy      492 -STAGGVENGCT-DRYFPALSOETGETLMQARLATVATGA--ISYELDGVOYIAGAG 547
Db      522 ISTAGNVLFIGATADNYLRAVYMSNGEKLWEARLP--AGGQATPMTYEVNGKQYVVISAG 579
Qy      548 GL-TYGTQLNAPLAEALIDSTSVGNAIYVVALP 578
Db      580 GHGSFGTKM-----GDYIVAVALP 598

```

Search completed: November 18, 2004, 05:21:08
 Job time : 45 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2004, 05:13:29 ; Search time 145 Seconds

(without alignments)
1414.068 Million cell updates/sec

Title: US-10-802-682-8

Perfect score: 3069

Sequence: 1 NMPTLLRTSAVILLTAPLAP.....AENIDSTGVGNALYFALPQ 579

Scoring table: BLOSUM62

Searched: 1570615 segs, 354127592 residues

Total number of hits satisfying chosen parameters: 1570615

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2765	90.1	580	14 US-10-162-713-8	Sequence 8, Appli
2	2569.5	83.7	578	14 US-10-162-713-5	Sequence 5, Appli
3	2559.5	83.4	579	14 US-10-162-713-7	Sequence 7, Appli
4	755.5	24.6	717	14 US-10-369-493-10840	Sequence 10840, A
5	624.5	20.3	567	14 US-10-369-493-7364	Sequence 7364, Ap
6	622	20.3	577	14 US-10-369-493-21739	Sequence 21739, A
7	620	20.2	692	14 US-10-369-493-20524	Sequence 20524, A
8	619	20.2	559	14 US-10-369-493-4607	Sequence 4607, Ap
9	578	18.8	623	14 US-10-369-493-8046	Sequence 8046, Ap
10	554.5	18.1	601	14 US-10-369-493-4930	Sequence 4930, Ap
11	547	17.8	592	14 US-10-369-493-7668	Sequence 7668, Ap
12	529.5	17.3	532	14 US-10-369-493-7687	Sequence 7687, Ap
13	524.5	17.1	706	14 US-10-369-493-10838	Sequence 10838, A

14	513.5	16.7	683	14 US-10-369-493-4929	Sequence 4929, Ap
15	482.5	15.7	592	14 US-10-369-493-15969	Sequence 15969, A
16	482.5	15.7	592	14 US-10-369-493-16311	Sequence 16311, A
17	482.5	15.7	593	14 US-10-369-493-15593	Sequence 15593, A
18	416	13.6	792	14 US-10-369-493-4878	Sequence 4878, Ap
19	412	13.4	803	14 US-10-369-493-7637	Sequence 7637, Ap
20	408	13.3	804	14 US-09-927-827-69	Sequence 69, Appli
21	408	13.3	804	14 US-10-369-493-4821	Sequence 4821, Ap
22	408	13.3	804	14 US-10-369-493-7580	Sequence 7580, Ap
23	407	13.3	798	14 US-10-369-493-15413	Sequence 15413, A
24	407	13.3	798	14 US-10-369-493-15781	Sequence 15781, A
25	407	13.3	798	14 US-10-369-493-16183	Sequence 16183, A
26	399	13.0	796	14 US-10-369-493-703	Sequence 703, Appl
27	373.5	12.2	777	14 US-10-369-493-8038	Sequence 8038, Ap
28	367.5	12.0	778	14 US-10-369-493-15182	Sequence 15182, A
29	363	11.8	786	14 US-10-369-493-16335	Sequence 16335, A
30	363	11.8	791	14 US-10-369-493-15596	Sequence 15596, A
31	363	11.8	803	14 US-10-369-493-15973	Sequence 15973, A
32	355	11.6	785	14 US-10-369-493-11883	Sequence 11883, A
33	354	11.5	783	14 US-10-369-493-14105	Sequence 14105, A
34	351.5	11.5	800	14 US-10-369-493-14064	Sequence 14064, A
35	342.5	11.2	783	14 US-10-369-493-11386	Sequence 11386, A
36	342.5	11.2	783	14 US-10-369-493-14430	Sequence 14430, A
37	342.5	11.2	783	14 US-10-369-493-14704	Sequence 14704, A
38	342.5	11.2	783	14 US-10-369-493-15187	Sequence 15187, A
39	324.5	10.6	788	14 US-10-369-493-11652	Sequence 11652, A
40	324.5	10.6	788	14 US-10-369-493-14639	Sequence 14639, A
41	323	10.5	787	14 US-10-369-493-14374	Sequence 14374, A
42	311	10.1	790	14 US-10-369-493-16392	Sequence 16392, A
43	176.5	5.8	564	14 US-10-369-493-17904	Sequence 17904, A
44	175.5	5.7	381	15 US-10-282-122A-50141	Sequence 50141, A
45	157.5	5.1	355	15 US-10-282-122A-69695	Sequence 69695, A

ALIGNMENTS

RESULT 1
US-10-162-713-8
; Sequence 8, Application US/10162713
; Publication No. US20030228672A1
; GENERAL INFORMATION:
; APPLICANT: Choi, Eun-Sung
; APPLICANT: D'Elia, John
; APPLICANT: Kim, Hye-Sun
; APPLICANT: Kim, Xi-Soo
; APPLICANT: Lee, Jung Kee
; APPLICANT: Pan, Jae-Gu
; APPLICANT: Stoddard, Steven F.
; APPLICANT: Yum, Do-Young
; TITLE OF INVENTION: Sorbicol Dehydrogenases of Ketogulonigenium spp., Genes and Method
; FILE REFERENCE: 1533.3740000
; CURRENT APPLICATION NUMBER: US/10/162.713
; CURRENT FILING DATE: 2002-06-06
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 8
; LENGTH: 580
; TYPE: PRT
; ORGANISM: Ketogulonigenium sp.
; US-10-162-713-8

Query Match 90.1%; Score 2765; DB 14; Length 580;
Best Local Similarity 87.6%; Pred. No 1e-236;
Matches 507; Conservative 38; Mismatches 34; Indels 0; Gaps 0;

QY 1 NMPTLLRTSAVILLTAPAAFAQVPTITDELLANPPAGGWINYGRQENYRHSPLTQIT 60
DB 1 NMPTLLRTSAVILLTGP1PAFAQVPTITDELLANPPAGGWINYGRQENYRHSPLTQIT 60
QY 61 ADNVOGLDVAWAGAGAGAGVGTPTMHDGWNVYLANPGVIGCALDAQCGDILLWERRLOPA 120

Db 61 TDVVGQQLVWARGMEAGAVQVTPMIHGVMYLANPGDVIQAIIDAKTGDLMEHRQLP 120
Qy 121 VATTNAOGDRKRGVALYGTSLYFSSWDNHLIALDMETGQVVPVERSGEDGLTSNTGP 180
Db 121 TSTLSSGDRKRGVALYGTSLYFSSWDNHLIALDMETGQVVPVERSGEDGLTSNTGP 180
Qy 121 VASLNGDRKRGVALYGTSLYFSSWDNHLIALDMETGQVVPVERSGEDGLTSNTGP 180
Db 181 IVANGVIVAGSTCOYSPYCGFISGHDSATGEBELMRNFIFIQAGEEDETMGNDPEARMT 240
Qy 181 IVANGVIVAGSTCOYSPYCGFISGHDSATGEBELMRNFIFIQAGEEDETMGNDPEARMT 240
Db 181 IVANGVIVAGSTCOYSPYCGFISGHDSATGEBELMRNFIFIQAGEEDETMGNDPEARMT 240
Qy 241 GVMGQITYPDTNLVYFVGSTGVGPASSETGPGTGLYGTNTREFAVRPDTGEIWMHQTL 300
Db 241 GVMGQITYPDTNLVYFVGSTGVGPASSETGPGTGLYGTNTREFAVRPDTGEIWMHQTL 300
Qy 301 PRDWMDECTEFEMVAVADVQPSAEMGLRINPNAATGERRVLTVGAPCKTGTMSFDA 360
Db 301 PRDWMDECTEFEMVAVADVQPSAEMGLRINPNAATGERRVLTVGAPCKTGTMSFDA 360
Qy 361 SGEFLMARDNTYMTASIDETGLVYVNEBDAVLKELDEVYDVCPTFLGGRDMSAALND 420
Db 361 TGEFLMARDNTYMTASIDETGLVYVNEBDAVLKELDEVYDVCPTFLGGRDMSAALND 420
Qy 421 TGIYFLPLNNAACVDINAADQEFSAIDVYNTSATAKLAPGENMGRIDAIDISTGRTLMSA 480
Db 421 TGIYFLPLNNAACVDINAADQEFSAIDVYNTSATAKLAPGENMGRIDAIDISTGRTLMSA 480
Qy 481 ERPAANSPLYSTAGGVFNGGTRFRALSOETGETLMOARLATVATGQALSYELDGVQ 540
Db 481 ERPAANSPLYSTAGGVFNGGTRFRALSOETGETLMOARLATVATGQALSYELDGVQ 540
Qy 541 YVIAAGGLTYGTQLNAPLAELIDSTSVGNAYVFPALPQ 579
Db 541 YVIAAGGLTYGTQLNAPLAELIDSTSVGNAYVFPALPQ 579

RESULT 2

US-10-162-713-5
; Sequence 5, Application US/10162713
; Publication No. US20030228672A1
; GENERAL INFORMATION:
; APPLICANT: Choi, Eun-Sung
; APPLICANT: D'Elia, John
; APPLICANT: Kim, Hye-Sun
; APPLICANT: Kim, Mi-Soo
; APPLICANT: Lee, Jung Kee
; APPLICANT: Pan, Jae-Gu
; APPLICANT: Stoddard, Steven F.
; APPLICANT: Yum, Do-Young
; TITLE OF INVENTION: Sorbitol Dehydrogenases of Ketogulonigenium spp., Genes and Method
; FILE REFERENCE: 1533.3740000
; CURRENT APPLICATION NUMBER: US/10/162,713
; CURRENT FILING DATE: 2002-06-06
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 578
; TYPE: PRP
; ORGANISM: Ketogulonigenium sp.
US-10-162-713-5

Query Match 83.7%; Score 2569.5; DB 14; Length 578;
Best Local Similarity 80.1%; Pred. No. 2,4e-219;
Matches 466; Conservative 59; Mismatches 55; Indels 1; Gaps 1;
Qy 1 MNPTTLRTSAVLLTAPAAFAQVTPITDELANPPAGEMINYGNOENYRHSPLTQT 60
Db 1 MKNSLILASVAVAFPAFAADVTPTVDELANPPAGEMINYGNOENYRHSPLNQT 60
Qy 61 ADVNGQLVWARGMEAGAVQVTPMIHGVMYLANPGDVIQAIIDAKTGDLMEHRQLP 120
Db 61 PDVNGQLVWARGMEAGAVQVTPLIHGVMYLANPGDIIQAIIDAKTGDLMEHRQLP 120

Qy 121 VATTNAOGDRKRGVALYGTSLYFSSWDNHLIALDMETGQVVPVERSGEDGLTSNTGP 180
Db 121 TSTLSSGDRKRGVALYGTSLYFSSWDNHLIALDMETGQVVPVERSGEDGLTSNTGP 179
Qy 181 IVANGVIVAGSTCOYSPYCGFISGHDSATGEBELMRNFIFIQAGEEDETMGNDPEARMT 240
Db 181 IVANGVIVAGSTCOYSPYCGFISGHDSATGEBELMRNFIFIQAGEEDETMGNDPEARMT 239
Qy 241 GVMGQITYPDTNLVYFVGSTGVGPASSETGPGTGLYGTNTREFAVRPDTGEIWMHQTL 300
Db 240 GVMGQITYPDTNLVYFVGSTGVGPASSETGPGTGLYGTNTREFAVRPDTGEIWMHQTL 299
Qy 301 PRDWMDECTEFEMVAVADVQPSAEMGLRINPNAATGERRVLTVGAPCKTGTMSFDA 360
Db 300 PRDWMDECTEFEMVAVADVQPSAEMGLRINPNAATGERRVLTVGAPCKTGTMSFDA 359
Qy 361 SGEFLMARDNTYMTASIDETGLVYVNEBDAVLKELDEVYDVCPTFLGGRDMSAALND 420
Db 360 TGEFLMARDNTYMTASIDETGLVYVNEBDAVLKELDEVYDVCPTFLGGRDMSAALND 419
Qy 421 TGIYFLPLNNAACVDINAADQEFSAIDVYNTSATAKLAPGENMGRIDAIDISTGRTLMSA 480
Db 420 TGIYFLPLNNAACVDINAADQEFSAIDVYNTSATAKLAPGENMGRIDAIDISTGRTLMSA 479
Qy 481 ERPAANSPLYSTAGGVFNGGTRFRALSOETGETLMOARLATVATGQALSYELDGVQ 540
Db 480 ERPAANSPLYSTAGGVFNGGTRFRALSOETGETLMOARLATVATGQALSYELDGVQ 539
Qy 541 YVIAAGGLTYGTQLNAPLAELIDSTSVGNAYVFPALPQ 579
Db 540 YVIAAGGLTYGTQLNAPLAELIDSTSVGNAYVFPALPQ 578

RESULT 3

US-10-162-713-7
; Sequence 7, Application US/10162713
; Publication No. US20030228672A1
; GENERAL INFORMATION:
; APPLICANT: Choi, Eun-Sung
; APPLICANT: D'Elia, John
; APPLICANT: Kim, Hye-Sun
; APPLICANT: Kim, Mi-Soo
; APPLICANT: Lee, Jung Kee
; APPLICANT: Pan, Jae-Gu
; APPLICANT: Stoddard, Steven F.
; APPLICANT: Yum, Do-Young
; TITLE OF INVENTION: Sorbitol Dehydrogenases of Ketogulonigenium spp., Genes and Method
; FILE REFERENCE: 1533.3740000
; CURRENT APPLICATION NUMBER: US/10/162,713
; CURRENT FILING DATE: 2002-06-06
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 579
; TYPE: PRP
; ORGANISM: Ketogulonigenium sp.
US-10-162-713-7

Query Match 83.4%; Score 2559.5; DB 14; Length 579;
Best Local Similarity 80.5%; Pred. No. 1.9e-218;
Matches 466; Conservative 52; Mismatches 60; Indels 1; Gaps 1;
Qy 1 MNPTTLRTSAVLLTAPAAFAQVTPITDELANPPAGEMINYGNOENYRHSPLTQT 60
Db 1 MKTSFLFAGVAAALASGTIALADVTPTVDELANPPAGEMINYGNOENYRHSPLNQT 60
Qy 61 ADVNGQLVWARGMEAGAVQVTPMIHGVMYLANPGDVIQAIIDAKTGDLMEHRQLP 120
Db 61 PENVGQLVWARGMEAGAVQVTPLIHGVMYLANPGDIIQAIIDAKTGDLMEHRQLP 120
Qy 121 VATTNAOGDRKRGVALYGTSLYFSSWDNHLIALDMETGQVVPVERSGEDGLTSNTGP 180
Db 121 VATTNAOGDRKRGVALYGTSLYFSSWDNHLIALDMETGQVVPVERSGEDGLTSNTGP 180

Db 121 VALLNSGEBIRGIALYGTINVFVSMNDNLVALDAATGQVTFDVEDGQED -MVSNSGP 179

Qy 181 IVANGVIVASTCOYSPYSGCFISGHSATGSELNRHFIPOGEEGDETWGNDPEARMWT 240

Db 180 IVANGVIVASTCOYSPYSGCFISGHSATGSELNRHFIPOGEEGDETWGNDPEARMWT 239

Qy 241 GWMGQITVDVNTLVFYSGTGVGPASSETQGTGPGTLYGNTNFAVRPDTGEIYVRHQTLL 300

Db 240 GWMGQITVDVNTLVFYSGTGVGPASSETQGTGPGTLYGNTNFAVRPDTGEIYVRHQTLL 299

Qy 301 PRODMOECTFEMVNAVVDYOPSAEMEGRAINPNAATGERRLTGAPCKGTMTMSFDDA 360

Db 300 PRODMOECTFEMVNAVVDYOPSAEMEGRAINPNAATGERRLTGAPCKGTMTMSFDDA 359

Qy 361 SGFFLARDNTYNTMLASIDETGLVTNEDAVLKEDEVEDVCPFTLGGSDWSAALNPD 420

Db 360 TGEFLMARDNTYNTMLASIDETGLVTNEDAVLKEDEVEDVCPFTLGGSDWSAALNPD 419

Qy 421 TGYFLPLNACADLAVIDEFTAMDYNTSATYLLAPEKENGRIDALDISGRTLMNV 480

Db 420 SGYFLPLNACADLAVIDEFTAMDYNTSATYLLAPEKENGRIDALDISGRTLMNV 479

Qy 481 ERPAANSPLSTAGGVFNCGTDRYFALSOETGETLMQARLATATVATGOAISYEIDGVQ 540

Db 480 ERPAANSPLSTAGGVFNCGTDRYFALSOETGETLMQARLATATVATGOAISYEIDGVQ 539

Qy 541 YIAIGAGGLTYGTQLNAPLAELIDSTSVGNATYFALPQ 579

Db 540 YIAIGAGGLTYGTQLNAPLAELIDSTSVGNATYFALPQ 578

RESULT 4

US-10-369-493-10840

/ Sequence 10840, Application US/10369493

/ Publication No. US20030233675A1

/ GENERAL INFORMATION:

/ APPLICANT: Cao, Yongwei

/ APPLICANT: Hinkle, Gregory J.

/ APPLICANT: Slater, Steven C.

/ APPLICANT: Goldman, Barry S.

/ APPLICANT: Chen, Xianfeng

/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

/ TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

/ FILE REFERENCE: 38-10(52052)B

/ CURRENT APPLICATION NUMBER: US/10/369,493

/ CURRENT FILING DATE: 2003-02-28

/ PRIOR APPLICATION NUMBER: US 60/360,039

/ PRIOR FILING DATE: 2002-02-21

/ NUMBER OF SEQ ID NOS: 47374

/ SEQ ID NO 10840

/ LENGTH: 717

/ TYPE: PRT

/ ORGANISM: Sphingomonas aromaticivorans

/ FEATURE:

/ NAME/KEY: unsure

/ LOCATION: (1) (717)

/ OTHER INFORMATION: unsure at all Xaa locations

US-10-369-493-10840

Db 153 VFTGLDRLVALDQKTKGVVMSKVVNPQEDYITGAPRVVKGKVLIGSGGSSEYKAG- 211

Qy 201 FTSCHDSATGSELNRHFIPOGEEGDETWGNDPEARMWTG---VMGQIT 247

Db 212 YIAADVNTGNEVVKFHFVPGNPADGFENKAMENARWAGEW---WKLGGGGTYMDSIT 268

Qy 248 YDPVNTLVFYSGTGVGPASSETQGTGPGTLYGNTNFAVRPDTGEIYVRHQTLLPDMWDQ 307

Db 269 YDPVNTLVFYSGTGVGPASSETQGTGPGTLYGNTNFAVRPDTGEIYVRHQTLLPDMWDQ 327

Qy 308 ECTFEMNAVNDYOPSAEMEGRAINPNAATGERRLTGAPCKGTMTMSFDDAASGEFLM 366

Db 308 ECTFEMNAVNDYOPSAEMEGRAINPNAATGERRLTGAPCKGTMTMSFDDAASGEFLM 365

Qy 367 ARDTYNTMLASID-ETGLVTNEDAVLKEDEVEDVCPFTLGGSDWSAALNPDGTLYF 425

Db 370 ATFPVNAVMTAGIDBKTKATVNPEARYEKGFVSLPGAVGAHSMQPOFSFKTGLLY 429

Qy 426 LPNNACT-----DINAVDOEF-SALDYNTS--ATKX--AAGFEMGRIDALDISG 474

Db 420 LPNNACT-----DINAVDOEF-SALDYNTS--ATKX--AAGFEMGRIDALDISG 473

Qy 475 RTMGAERPAANSPLSTAGGVFNCGTDRYFALSOETGETLMQARLATATVATGOAISY 534

Db 475 RTMGAERPAANSPLSTAGGVFNCGTDRYFALSOETGETLMQARLATATVATGOAISY 533

Qy 535 ELDGVQYTA--IGAGGL 549

Db 550 ELDGVQYTA--IGAGGL 548

RESULT 5

US-10-369-493-7364

/ Sequence 7364, Application US/10369493

/ Publication No. US20030233675A1

/ GENERAL INFORMATION:

/ APPLICANT: Cao, Yongwei

/ APPLICANT: Hinkle, Gregory J.

/ APPLICANT: Slater, Steven C.

/ APPLICANT: Goldman, Barry S.

/ APPLICANT: Chen, Xianfeng

/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

/ TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

/ FILE REFERENCE: 38-10(52052)B

/ CURRENT APPLICATION NUMBER: US/10/369,493

/ CURRENT FILING DATE: 2003-02-28

/ PRIOR APPLICATION NUMBER: US 60/360,039

/ PRIOR FILING DATE: 2002-02-21

/ NUMBER OF SEQ ID NOS: 47374

/ SEQ ID NO 7364

/ LENGTH: 567

/ TYPE: PRT

/ ORGANISM: Burkholderia cepacia

US-10-369-493-7364

Query Match 24.6%; Score 755.5; DB 14; Length 717;

Best Local Similarity 33.6%; Pred. No. 7.7e-58;

Matches 187; Conservative 90; Mismatches 217; Indels 63; Gaps 17;

Db 28 ITDELLANPPAGEINYGNGQENYRHSPLQITADNVGQOLVW-----ARGMEGAVQ 81

Qy 38 VTDALINQAEKGELSGRDYGEORSPLOINDANGVGLMFFDLRTARQEA----- 93

Db 82 VTPMHDGVNVLNPGDVLALDQGTDLMEHRRQLPAVALTNAQDR-KRGVALYGTG 140

Qy 94 -TFLMHDGTLVYISAWGNVAFADAKTGALKMSYDPEVRETLVACODAVNAGVALYDK 152

Db 141 LYSSSWNHLLALDMEGGVFVDERSGSDGLTSNTGTIVANGVIVASTCOYSPYGC 200

Query Match 20.3%; Score 624.5; DB 14; Length 567;

Best Local Similarity 28.5%; Pred. No. 2.4e-46;

Matches 172; Conservative 105; Mismatches 257; Indels 69; Gaps 21;

Db 3 PTTL-LRTSAVLLITAPAFAPQVPT--DELIANPPAGWINYGRNQENYRHSPLTQI 59

Qy 6 PTTL-LRTSAVLLITAPAFAPQVPT--DELIANPPAGWINYGRNQENYRHSPLTQI 58

Db 66 DTANVAKLQVMSYKFPADLQGFBA-----TPVNGRYLVTTPKKNVYAFDAATGQL 120

Qy 112 WEHRQLPAVALTNAQGD-RRGVALYGTSLYSSWMDNHLLALDMETGQVFDVERSGGE 170

Db 121 WKFEPKLAESFKTACCBVIRNGVALYKNTYVAMLGSDVVALDQGLALW--RKQWFE 178

Qy 171 DGL-TSNTGTIVANGVIVAST-CQYSPYGCFSGHSATGSELNRHFIPOGEEGDE 228

Db 179 PGLGVAFLSLPLALDGLVVGSAAGGYGARG-FIALALNDPNGNVLWKFTVPAGEKAD 237

Qy 229 TWGNDEAR-----NMGTGVMGQITTDPTNLVFGSTGVGPASETORGTGGTLTGNTNR 283

Db 238 TWNGNGEHBGAPAMLTG-----TYDASKTLYWGVNGPMLADLR--PGNNLY-SDSL 289

Qy 284 FAVRPDTGELVTHQTLPRDNWDOCTEFEMMVANVDVQSAEMEGRAINPMAATGERV 343

Db 290 LADPFTGDKMHYQYTKHDITWDYDVTNVLATITTYQ-DKEVDALIHADR----- 340

Qy 344 LTGAPCKTGMWSEFDSAAGEFLWARDNTNMIASIDETGLVY---VNEDAVLKELDVE 399

Db 341 -----GFFHALDGTGGLYAKP-----FTATSVYGYTADGAPIDQAKYPRAGT 387

Qy 400 YDVCPFTLGRDMSSAALNDPTGIYPLPLNNACYDIMAVDOEF-SALDYNTNTAKLAP 458

Db 388 IETCPSTLGGKMWMSISYDPDKHIAVVPALHACMSISGKSVNMEGLPLIGRGFEIKPRP 447

Qy 459 GFENMGRIADISTGRTLMSARPAANYSPVLSTAGVVPNGCTDRFYRALSQETGETL 518

Db 448 GSKYGELOAINVDITGKWSHMSKLPWNGVATTGGLAFSGSLDGLHYADPTTGTLY 507

Qy 519 WCA-RLATVATGQAISYELDVQYIAGAGLTYGTQ---LNAFLAELIDSTSVGNALYV 574

Db 508 WQBPCLASGIVAQPSYFEVVDGKEYVALAG---YGGANPIWGGPMAKAEKVPBGCTLYV 564

Qy 575 PAL 577

Db 565 PAL 567

RESULT 6

US-10-369-493-21739

Sequence 21739, Application US/10369493

Publication No. US2003023675A1

GENERAL INFORMATION:

APPLICANT: Cao, Yongwei

APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.

APPLICANT: Goldman, Barry S.

APPLICANT: Chen, Xianfeng

TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

FILE REFERENCE: 38-10(52052)B

CURRENT APPLICATION NUMBER: US/10/369,493

CURRENT FILING DATE: 2003-02-28

PRIOR APPLICATION NUMBER: US 60/360,039

PRIOR FILING DATE: 2002-02-21

NUMBER OF SEQ ID NOS: 47374

SEQ ID NO 21739

LENGTH: 577

TYPE: PRT

ORGANISM: Rhodobacter capsulatus

US-10-369-493-21739

Query Match 20.3%; Score 622, DB 14; Length 577;

Best Local Similarity 29.6%; Pred. No. 4.2e-46;

Matches 186; Conservative 98; Mismatches 231; Indels 114; Gaps 28;

Qy 11 AAVLLTPAPAFQVPTDELAN--PPAGEWINGRQENYRHSPLTOITADNQGLO 68

Db 2 AAVALTLC-FGTADAGVTEDLRNDQGTAGDVNTGMSGDLORFPLATLNKRDVKNLM 60

Qy 69 LWA-----RGEAGAVQVTPMTHDGVNLANPQVDVIALDAQTGLIWEHRQLPA 120

Db 61 PAWAFSLGSEKRGQES-----OPIVYDGMVITGSYSRLYALDITGKEIMWQYDARLE 115

Qy 121 VATLNAGDKRKGVALVGTSLYSSWMDNLIALDMETGQVFPVERSGSGEDGLTSNTGP 180

Db 116 -GLIPCCDVNRGAALIGDVTFTGLDARIVALDRKTDGVKMKKADYKEGI-STTAAP 173

Qy 181 IVANGVIVAGSTQYSPYCGF--ISGHSATGSELWRNHFIPQSGEG----- 226

Db 174 LTVNGLVITGNS--GSEFQIVGEVQANNAETGELWTRPMI--EGHMTLNGKSPMTGT 229

Qy 227 -DETWGNDFEARWMTG---VWQITTDPTNLVFGSTGVGPASETOR--GTP-----GG 275

Db 230 LNAITPDDL---WKTGGGATTLGSSYDAETDLTVFGAGNAPWNSHLRNAGTEBGNKGD 286

Qy 276 TLVGTNREFNRPDTGIVRHOFLPRDNWDOCTEFEMMVANVDVQSAEMEGRAINPN 335

Db 287 NLVYAA-SRVGIDPFTGILKMHFQITTPREGWDFGVNE-VGFPD-----KDG---NKR 334

Qy 336 AATGEREVLTCAPCKTGMWSFPAASGEFL--WARDNTNMIASIDETGLTVNED--- 390

Db 335 FATADR-----NGFYVLNRADGKFVAMPFVNIT--WAKGIDETGRPIYNEENRP 384

Qy 391 -----AVLKELDVEDYDCTFELGRDMSAALNDPTGIYPLPLNNACYDIMAVDOERSAL 445

Db 385 GAPPAALAAKQGITFAVPSFLGGKMWPMAYSCQTLTFYVPSNE-----WGM 442

Qy 446 DVYN-----TSATKLAPG-----ENMGRIADISTGRTLMSARPAANYSPYLST 493

Db 433 DINNEPINKKGAAYLAGFTIKPLFEDYIGSLKADPNTGAVKMEKNDAPLMGVMTT 492

Qy 494 AGSVFENGCTDRFYRALSQETGETLQARLATVATGQAISYELDVQYIAGAGLT- 550

Db 493 AGKLVFTGNPEGEFWALDAETGKLMSPQTSGSIVGQPIWEDDGEQVSIISGWSGAVP 552

Qy 551 -YGTQNLNAPLAELIDSTSVGNALYVFPALP 578

Db 553 LWGE-----VAKRNYNLNQGCTVWTFRLP 577

RESULT 7

US-10-369-493-20524

Sequence 20524, Application US/10369493

Publication No. US2003023675A1

GENERAL INFORMATION:

APPLICANT: Cao, Yongwei

APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.

APPLICANT: Goldman, Barry S.

APPLICANT: Chen, Xianfeng

TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

FILE REFERENCE: 38-10(52052)B

CURRENT APPLICATION NUMBER: US/10/369,493

CURRENT FILING DATE: 2003-02-28

PRIOR APPLICATION NUMBER: US 60/360,039

PRIOR FILING DATE: 2002-02-21

NUMBER OF SEQ ID NOS: 47374

SEQ ID NO 20524

LENGTH: 692

TYPE: PRT

ORGANISM: Rhodospseudomonas palustris

US-10-369-493-20524

Query Match 20.2%; Score 620, DB 14; Length 692;

Best Local Similarity 29.3%; Pred. No. 8.3e-46;

Matches 180; Conservative 100; Mismatches 246; Indels 88; Gaps 23;

Qy 15 LITAPAAFAQ-----VTPITDE--LLAN-PPAGEWINGRQENYRHSPLTOIT 60

Db 5 LITPAVAAQDSKSAHIRAVTGAVDSAAIVANVTKTKMPSGGLDYAFERFSKIDQIN 64

Qy 61 ADNVGQLOLVN--ARGMEAGAVQVTPMTHDGVNLANPQVDVIALDAQTGLIWEHRQL 118

Db 65 TENKQGLQLOWSYLSGSERG--VEATPVAVDGMIVTWSGVSVAHALDTRIKKLTPEPKV 123

Qy 119 PAVATLNAQGR-----KRGVALYGTSLYSSWMDNLIALDMETGQVFPVERSGSGEDG 172

Db 124 D-----HSKYGRCGCDVNRGVALYKGVFVGAVDGLIALDAATGSKAMEIDTLIDHEH 178

Qy 173 LFSNTTGPVANGVIVAGS--TCQYSPYCGFISGHSATGSELWRNHFIP-----QP 222

Db 179 SYITGAPRVENGKVIINGCAEYGAR-G-VTAAYDAETGKQAMRFTVPDPSKEFEDES 237
 QY 223 GEGDETFMGNDPEARWMTG-----VMGQITVDPVTNLVFGSTGVPASETORGTGPGTILY 278
 Db 238 MEKRAKTM-DPAKMWMLNGGGGTAMOTITPDPDLNLIYGTGNGSPMNHLSRPAAGDNL 296
 QY 279 GTMTRFVRPDTGGEIYWRHQTLPREDWDECTEFEMVAN--VDVQSAEMEGIRAINPNA 336
 Db 297 YLASIYALNADTKRYWHYQETPGDWDYSTOPMLADIAIDKPE----- 342
 QY 337 ATGERRLTGAPCKGTGTMGFFDASGEFLWADNTNTNMIASIDETGL-VTVNEDAVLKE 395
 Db 343 ----RKYIHLAP-KNGFFVIDRTDGKFTISAKNFVDVMAATGDSNGRPEVE---AS 394
 QY 396 LDVEYDVCTPFLGDRWSSAALNPDTGIYFLPLNNAACYD--MAVDQ-----EFSALD 446
 Db 395 ADSKFPDIPGPYGAHMHHPMSFNPQTGLVYLPAQGVVNLGKALTONMEPEFKGSGTT 454
 QY 447 VYNTSANTAKLAPGEMN--GRIDAIDISTGRTLSAERPAANYSPVLSAGVVFNGGTD 504
 Db 455 GKNVGFETLNAVPP-KNLPFGLVAMDPVOQKEVWRLEIYSPNNGTTLITRAGNLVFGCTAD 513
 QY 505 RYFRALSOETGETLMQARLATVATGCAISYELDVOYIAGAG-GUTYGTOLNAPLAEAI 563
 Db 514 GRFVAVYAKTGKELMESPLGTGAVAAPATYWDVOYVIAVGMGVFGISARATETEA- 572
 QY 564 DSTSVGNATVFPAL 577
 Db 573 ----PGTVYTFVAV 581

RESULT 8

US-10-369-493-4607
 ; Sequence 4607, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; PRIOR FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 4607
 ; LENGTH: 559
 ; TYPE: PRT
 ; ORGANISM: Burkholderia fungorum
 US-10-369-493-4607

Query Match 20.2%, Score 619, DB 14, Length 559,
 Best Local Similarity 28.3%, Pred. No. 7,4e-46,
 Matches 169; Conservative 105; Mismatches 256; Indels 68; Gaps 20;
 QY 7 LRTSAATLLTAPAFQAVTPI--DELLANPAGEWINVRGNOENYRHSPLQITADYV 64
 Db 3 MSVAVASLVFGSAADDPATYERLTAAQSDPGMLTYRYYNGOAHSEPLQIOTANV 62
 QY 65 GQLOLVNA-----RGMEAGAVQVTPIHIDGVVYLANPBGDIVIQAIDAGTGLIWEHR 116
 Db 63 KNLQVMSYKFPADLQGGFEA-----TPIVNGRYLFTVTPKDNVYAFADATGKQWFEF 117
 QY 117 QLPRAVATLANQGD-RKRGVALLYGTSLYFSSMDNHLILDMETQGVVDVVRSGSGEGL- 174
 Db 118 KLGESSEKYLACCDVINGVALYGGKVVVAMSGDVVALDQGTGLAW--RKQNFPEGLG 175
 QY 175 SNTGTPIVANGVIVAGST-CQSPYGCFIGSHDSATGEELRNHFTLPQGEESDETGWND 233

Db 176 AFSLAPLADGALVYGAAGEYGAR-G-FIALANPDNGVIMKRTVPAPAGEKADTPWNG 234
 QY 234 FEAR-----NMWGWGQITVDPVTNLVFGSTGVPASETORGTGPGTILYNTFRVAPR 288
 Db 235 MOEHOGAPAMLT-----TYDAASKTLYWGVGNPGLADLR--FGDNLV-SDSILALDF 286
 QY 289 DTGEIYWHQQLPDPDNMQECTEFEMVANVDVQSAEMEGIRAINPNAATGERRLVIGAP 348
 Db 287 KTGDLMHYQYTKHDTPWDVNTPVLTATIRYQ-DKEVDALIHADR----- 332
 QY 349 CKTGTMGFFDASGEFLWADNTNTNMIASIDETGLVTV--VNEDAVLKELDVEYVCP 404
 Db 333 ---GEFHAIIDGTGLIYAKP-----PVLASVGYIADAGAPIDQASKYPRAGTTEFCP 384
 QY 405 TPLGDRWSSAALNPDTGIYFLPLNNAACYDINAVDQEF-SALDVYNTSATKLAPEEM 463
 Db 385 SFLGCKMWSISYDPDKIAVAVPALHACMSLSGKSVYMEGLPYLGGEFLKPEBPGSKGY 444
 QY 464 GRIDAIDISTGRTLSAERPAANYSPVLSAGVVFNGGTD RYFRALSOETGETLMQA-R 522
 Db 445 GELQALINDTGKRWSHMSKLPNNGVATTTNGLAFSGSDGHLIYAPDETGTGKYLQMSPK 504
 QY 523 LATVATGCAISYELDVOYIAGAGLTYGTQ---LNAPLAEADSTSVGNATVFPAL 577
 Db 505 LASGIVAPSVFEVDGKEYVAIAG---YCGANPIWGGPMAKAEKVPBGGLTVFPAL 559

RESULT 9

US-10-369-493-8046
 ; Sequence 8046, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; PRIOR FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 8046
 ; LENGTH: 623
 ; TYPE: PRT
 ; ORGANISM: Rhodobacter sphaeroides
 ; NAME/KEY: unsure
 ; LOCATION: (1) (623)
 ; OTHER INFORMATION: unsure at all Xaa locations
 US-10-369-493-8046

Query Match 18.8%, Score 578, DB 14, Length 623,
 Best Local Similarity 27.8%, Pred. No. 3,8e-42,
 Matches 177; Conservative 103; Mismatches 259; Indels 98; Gaps 22;
 QY 3 PTLTARTS-AAVLTLTAPAFQAVTPIDELL-ANPAGEWINVRGNOENYRHSPLQIT 60
 Db 10 PMKMLKTGIVATLILSSAPVFA-----NDSVLKATDDPKQALIQDGVANRYSELNQIT 64
 QY 61 ADNVGQLOLVNA-----RGMEAGAVQVTPIHIDGVVYLANP-BGVIOALD-AQTGDLIM 112
 Db 65 RENVEKLGQAVMTFSTGVLRGHEG-----SPLVVGIMYVHTPFPNNVVALDLANDGKILW 119
 QY 113 EHR-RQLPAVATLANQGRKRGVALLYGTSLYFSSMDNHLILDMETQGVVDVVRSGSGED 171
 Db 120 KYEPQNPDPVIGVMCCDTPVNRGVALYADGKIFLHQADITVVALDASKSVKRYTVQNGDPAK 179
 QY 172 GLTSNTGTPIVANGVIVAGSTCQSPYGCFIGSHDSATGEELRNHFT----- 219

[illegible]

```

RESULT 10
US-10-369-493-4930
? Sequence 4930, Application US/10369493
? Publication No. US20030233675A1
? GENERAL INFORMATION:
? APPLICANT: Cao, Yongwei
? APPLICANT: Hinkle, Gregory J.
? APPLICANT: Slater, Steven C.
? APPLICANT: Goldman, Barry S.
? APPLICANT: Chen, Xianfeng
? TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
? TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
? FILE REFERENCE: 38-10(52052)B
? CURRENT APPLICATION NUMBER: US/10/369,493
? CURRENT FILING DATE: 2003-02-28
? PRIOR APPLICATION NUMBER: US 60/360,039
? PRIOR FILING DATE: 2002-02-21
? NUMBER OF SEQ ID NOS: 47374
? SEQ ID NO 4930
? LENGTH: 601
? TYPE: PRT
? ORGANISM: Burkholderia fungorum
US-10-369-493-4930

```

Query Match	18.1%;	Score 554.5;	DB 14;	Length 601;
Best Local Similarity	26.7%;	Pred. No. 4.5e-40;		
Matches 169;	Conservative 111;	Mismatches 257;	Indels 97;	Gaps 26

```

QY 7 LRISAANILITTAAPAF-----AQVPTTBELLANPAGSAMIYKGNCHNOYHSRSPLOTAD 62
DQ 3 LRITTVGLAITSABALSSFVAQADSDQDGLMGNP--SMAAAQADYANHRISPLKQINEN 60
QY 63 NVGQLOLWARGNEGAV--QVTPMHDGWTYLANP-GDVIQALDAQDGLIWEH-RRQ 117
DQ 61 NVGKLGQVAMT--NSTVLKRGHGAPLVIGDWTYHSPPNNKYIATNLKQDQFIQWLPKQ 118
QY 118 LPAPVATLMAQGRKRGVALYGTSLYFSSNDNLILMNETQGVYFFDYERGSGBEDGLTSNT 177
DQ 119 DDQVSWMCCDITVNRGLAYGDGKIFLPQADPTLVLINAKTDGVYVMTAQNGNPKAGET-NT 177
QY 178 TGSIVANGVIVAG-STQCQSPYGCPIGSHDASTGBELARN----- 217

```

```

Dd 178 NAPHVGDGKVLTGISGGEVYAGRLIA-YDIKTKPAMTAVSTGPDKMDLIDDKUTTYA 236
Cy 218 ---FIPQPEBEDETKMGDNFEARMNTS---VMGQITIDPTNLNLYFSGVSGASETORG 271
Dd 237 DGMNVFVADGSSLSKSGKD---QMGLGGCTTGWYANDPKNLVLYYTGNGPITNPTOR- 292
Cy 272 TPGGTLYGNTRFAYVRPTGELVWHSQTLPRDNDQDCTEFEMVANYDVQ---PSA--- 324
Dd 293 -PDNNKMSMI-PARDLNTGQARWYVYQWTPRDEMDYGVNMILSLIDGKKVPAIVHF 350
Cy 325 EMEGFLAIPNPAATGERRVLTGAPCKTGTIMS--FDAASGEFLMARNTNTNMLASIDEI 382
Dd 351 DRNGF-GYTLINRETQO--LVAHQEPDPAVNADHVKMSGKPI--RNAASTQPAASDEN 405
Cy 383 GLVTNVEDAVLKELDVEVDVCTPFGGRDWSALINPTGTGFLPLNNAICYDIMAADOEE 442
Dd 406 -----VKGICPALGSDQOPAAVDSSSLFLVPTNHCMDVEPPDVY 449
Cy 443 SALDVYNTSATAKLAFG--FENMGRIDAIDISTGRTLSAERPANYSPVLTAGGVF 499
Dd 450 VSGQPY-VGATLSMTPGPRDNNMSMGNFTAMPDASGKIWMSKPERFSWGLTATAGVAF 508
Cy 500 NGGTDEYFPAALQETGETLMQARLATAVTATGALSYELDGVQYAL--GAGGLYGTQUNA 557
Dd 509 YGTLGEYIKAVRIKQGEKLMRPKTPSGIIGVFTFYERQKQFQVSGISGMA-GIGMA 567
Cy 558 PLAEIDS-----TSVGNALYFALP 578
Dd 568 GLEKSTEGAGVGVRELAKYTTALOCGLTFVPAIP 601

```

```

RESULT 11
US-10-369-493-7688
; Sequence 7688, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 7688
; LENGTH: 592
; TYPE: PRT
; ORGANISM: Burkholderia cepacia
US-10-369-493-7688

```

Query Match	17.8%;	Score 547;	DB 14;	Length 592;
Best Local Similarity	26.5%;	Pred. No. 2e-39;		
Matches 166;	Conservative 112;	Mismatches 254;	Indels 94;	Gaps 26;

```

OY      12  AVLLTAPAPAF-AOYTPITIDELIANPPAGEININGRQOEYRHSPLQIYIADNVGQLOLV 70
Dp      2  ALLSAAALSFLVAQADSDQDGLMKP--SNMAAQAGDYANHRYSPLKQIENNVIKLOVA 59
OY      71  WARGMEAGV--OVTPEITHGWTYLANE-GDVIQALDACTGDLINHE-RQOLPAVALTN 125
Dp      60  WT--MSTGVLRGHEGAPLVIGDWTYIHSPEFNKVIATINLDOFIWQYFKODQOVVSWA 117
OY      126  AQGDRKRGVALYGTSLYFSSWNHILADMETQGVFDERGSGEQLTNTTGPYVANG 185
Dp      118  CCDTVNRGLAYSDGKIFLQADDTKVLVALNAAATGCVVNTAONGPKAGEI-NINAPHVFGD 176
OY      186  VIVAG-STQOYSPYCQFISGHDSATGEELMNNH-----FIPQ 222

```

Db 177 KVLTSIGSGEFGVGRLLA-YDIKTKGPAWTAISTGDKMLIDPKITTYADGKRVPG 235
 Qy 223 GEEGDETWGNDPEARWMTG--VWGQITYPVTNLVFGSTGVGPASSETGRTGPGTLG 279
 Db 236 ADSLSKWKMD---QWMLGGGTTMGWYAMPKLNLYVYGTGNPGTNPTOR--FGNNKKS 290
 Qy 280 TNRFAVRPTGTGLVWRHQTLPNDNDQECTFEMMVANDVQ---PSA---EMEGRLAI 332
 Db 291 MSI-FARDLNTGQARWYQMTPHDEWDYDGVNEMILSDLSIDKKVPAIVHFRNGF-GY 348
 Qy 333 NPNAATGERVLTCAPCKTGTMWS--PDASGEFLMARDTNTNMIASIDETGLVTVNED 390
 Db 349 TLARETQ--LVAQKDPRAVNMADHYDMKSGPI--RNAAYSTQAGSHN----- 396
 Qy 391 AVLKEIDVEDVCPFLFGHDWSSALNPDTGIYFLPNNACYDIMAVDOESALDVNT 450
 Db 397 -----VKGICPAALGSKQCPAAVDPGSSILFLVPTNHVCMDEBFDVDYVSGQPY-V 447
 Qy 451 SATAKLAPG---FENMRIDAIDISGRTLSAERPAWSPYLSNAGVVFNGTDYF 507
 Db 448 GATLSMTPGNDNNSWGNFIAMDSKGIWMSKRPSPVMSGVLATAGVAFYGLGTY 507
 Qy 508 RALSOETGETLMQARLATVATGQAI SYELDGVQYIAI--GAGGLTYGTOLNAPLAEIDS 565
 Db 508 KAVRIKQKELMRKTPSGIIGNVFTYEQKQFIVYSGIGGWA-GIGWAAGLEKSTEG 566
 Qy 566 -----TSVGNATYFALP 578
 Db 567 LGAVGVRELAKYALGGLTFVFAIP 592

RESULT 12
 US-10-369-493-7687
 ; Sequence 7687, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 7687
 ; LENGTH: 532
 ; TYPE: PRT
 ; ORGANISM: Burkholderia cepacia
 ; US-10-369-493-7687

Query Match 17.3%; Score 529.5; DB 14; Length 532;
 Best Local Similarity 27.6%; Pred. No. 6.3e-38;
 Matches 161; Conservative 93; Mismatches 245; Indels 85; Gaps 20;
 Qy 27 PINDELL--ANPAGIEWINGRNOENRHSPLQITPDNNGOQ--LWARGHEAGAVQ 82
 Db 2 PVSQEQDRATSTNDLLNSGSAQTRYRGAQINKNTVSKLPAPITQTAANE-SMET 60
 Qy 83 TPMHDGVMYLANGVYQALDAQTDGLIWEHRROLPAVATLNAQGRKRGVALYSTLY 142
 Db 61 APTNSNVMFITTSFNNVAVDAVTKEFHGHKMGAVTF--CCGNNRGVALIADRLY 119
 Qy 143 FSSVDNHLIALDMETGCVFDERGSGEDGITSNTTGPYANGVYAGSTCYSPGCC-- 200
 Db 120 MGLTADALVALDAGTGVLMQTOIADPDEG-SETVAPTVADKVLGT--NGEGGIRG 176
 Qy 201 FISGHDATGEELRNHFIPQGEGBDETM-----GNDFEA 236

Db 177 FLKAFDANSGLLMTFTYIIPETQEG--VWATKDATGDKRIDDAEKKQIAEKGGDFK 234
 Qy 237 RMTGVMQCIYVDPTNLVFGSTGVGPASSETGRT--PGGLTYGNTNFRVAPDTEIV 294
 Db 235 TLGGGVWNAAPLDRQTHVVF---VGNPSPDLYGALIPGNLY-TDSLVAIDIDTGKX 290
 Qy 295 WRHQTLPNDNDQECTFEMMVANDVQPSAMEGLRAINPNAAGRGVYLTGAPCKITM 354
 Db 291 WHQYVPHVDVLDVDSPPMLIDV-----RDNNGMIRG---VHGG--KTGHV 334
 Qy 355 WSFDSAAGEFLMARDTNTNMIASIDETGLTVNEDAVLKEIDVEDVCPFLGGRDSS 414
 Db 335 YHDAIATGRLI-----RISQ--AMIPQNMNTLTPAAGARPL-----PGANGVEMSP 380
 Qy 415 AALNDPTGIYFLPNNACYDIMAVDOESALDVNTSATKLABFENMRIDAIDISTG 474
 Db 381 NAFDHTRLVYAANLHQMTYQVEDAAYPGGSKMLMGAFKTIASEQQMGKLSAVNDTG 440
 Qy 475 RLMSAERPAWSPYLSNAGVVFNGTDYFRLASGETGETLMQARLATVATGQAI SY 534
 Db 441 KVAADYKTEOPLIGGLVLTAGLVFNGGNGLFRAPDSATKRLWEPCCGAVNAPVSY 500
 Qy 535 ELIDGVQYIAGAGLTYGTOLNAPLAEALDSTVGNATYFALP 578
 Db 501 MVHGQYIAVAAAG--NTQLD-----FKGNVTVLFAIP 532

RESULT 13
 US-10-369-493-10838
 ; Sequence 10838, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 10838
 ; LENGTH: 706
 ; TYPE: PRT
 ; ORGANISM: Sphingomonas aromaticivorans
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(706)
 ; OTHER INFORMATION: unsure at all Xaa locations
 ; US-10-369-493-10838

Query Match 17.1%; Score 524.5; DB 14; Length 706;
 Best Local Similarity 29.0%; Pred. No. 2.7e-37;
 Matches 177; Conservative 79; Mismatches 244; Indels 111; Gaps 25;
 Qy 15 LITAPAAQVPTPINDELLANPAG--EWINNGRNOENRHSPLQITPDNNGOQOLV-- 71
 Db 22 IASPPAPSAANVLE-----GAGERPPGGGDKTHSRRLTRINANENVRGLAKOV 75
 Qy 72 ---ARGNEAGAVQYTPMHDGVMYLANGVYQALDAQTDGLIWEHRROLPAVATLNAQ 127
 Db 76 ELGTLRQGEA---TPVVVGVLYTSGTGRAVAFDAATGKELMRFBPEVDMQVNRVVC 130
 Qy 128 GPR-KRGVALIGTSYFSSVDNHLIALDMETGCVY---FVERGSGEDGLTSNTG--PI 181
 Db 131 CDVNRGVAAGVGRKVFVSLDGMMYALDARFGAAVWWSDFLENKAGG---NSTGAP 185
 Qy 182 VANGYIVAG-STCQYSPYCGFISGHDATGEELRNHFIPQGEGBDE-----TW 230

Db 186 IAGDVVVGKAGABYVRG-YTALDIDIGKLRNKHVVRPRDKLGPQETPELEALKTW 244
Qy 231 GNDPEARMWTG---VWGQITTYDPTVNLVFGSTGVGPASETORCGTGGTLVGTNTRFAY 286
Db 245 --DPNSKWDIGGGSPMDALINVDPEGLVVGNGGCPATSKRSPAGGNLVLASLVAL 302
Qy 287 RPTDGLVWHQTLPRNNMOECTPEMMVANDVQSAEMEGRLAINPMAATGERVYLVG 346
Db 303 DPKGRKRMHYQETPGGNMDFATQPMIFTRMKIDE-----DRPVLAH 346
Qy 347 APCGTGMWSPDASGEFLWARDNTYNNIASIDE-TGLVTVNEDAVLKELDEYD---- 401
Db 347 AP-KNGGLVILYDRDGLLAFANPIVRNMAKGDPTG-----RPLDPAADYTGPK 399
Qy 402 -VCPTEFGDMSSALNPTGIY-----FLP-----LNNACDIYMAVD 439
Db 400 IVPFATGANNHWPASDPATGLYGAVALDMGNLIMTPGAKPLKARGLNDDALLFTPD 459
Qy 440 --CEFSALDYNTSATVTKLAPGEN-----MGRIDALDISTGRTLSAERPA-ANYSPV 490
Db 460 VKELAAFPPEFGDAVKL-PAYQELAKOPATQOIRALIDPLTKITWADDTAGWQDRGV 518
Qy 491 LSTAGGVVNGGTDYFRALSOETETLMQARLATVATGQALSYELDVOYIAI---GA 546
Db 519 LTTSSGLTIGGYTGKLPVADTYTGKLIKIDTGTPIMAPMTYEVGVQYIAVWAGWG 578
Qy 547 GGLTYGTQOLNA 557
Db 579 GGYPFVPRISA 589

RESULT 14

US-10-369-493-4929
Sequence 4929, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 4929
LENGTH: 683
TYPE: PRT
ORGANISM: Burkholderia fungorum
US-10-369-493-4929

Query Match 16.7%; Score 513.5; DB 14; Length 683;
Best Local Similarity 26.9%; Pred. No. 2,4e-36;
Matches 158; Conservative 94; Mismatches 245; Indels 91; Gaps 20;

Qy 19 PAAFA-----CWTPTDEL--ANPPAGEWINGRNOENYHSPLOI 59
Db 9 PAAFAAEVONSSSTTSSPVSITQPVSGQOLDRATTSNDWLSNGSYAQRTFYGAQI 68
Qy 60 TADNVGLO--LVWARGEMAGAVQVTPMIDGVNMYLANPDVQIALDAQGLIWEHRQ 117
Db 69 NKTWVSLRPAFIPOZAVNE-SMETAIIVSGVWFITTSFNHYAVDAVATGKEFMYKAK 127
Qy 118 LPAVATLNAQGRKRGVALYGTSLYFSSWMDHLIALDMETGOVFPVERSGEGDGLTSMT 177
Db 128 MGAVTTF--CCGPNNRGAIAGDRLYMGTLDAKVALDAKTSVLMQTLADPREGT-SST 185
Qy 178 TGPVANGVIVAGSTCOYSPYGC--FISGHSATGEELWNRHPIPOGEEGDETW----- 230

Db 186 MAPTVVDGKVLIGT--NGEYIGRGLKAFDANSQOLMFWFYLLPENGQEG--VMATKDA 241
Qy 231 -----GNDPEARMWTGWGQITTYDPTVNLVFGSTGVGPASETORG 271
Db 242 TGRDGDKIDIAEKKQOLAEKGGDFYKTLGGGVWMAPIADROTHVTF--VWGNPSPLYG 288
Qy 272 T--PGTLYGNTRFAVRPTDGLVWHQTLPRNNMOECTPEMMVANDVQSAEMEGRL 329
Db 289 AIRPQDNLY-IDSLVALDIDLGKTKMHQYFPHDWDLDVSPMLIDV----- 346
Qy 330 RAINPNAATGERRVLTGAPCKTGTMWSPDASGEFLWARDNTYNNIASIDEFTGLTVNE 389
Db 347 RDNNGRMIPG---VWHG--KTGHVYVHDBRATGRLI-----RYSQ--AMTPOENMTLPT 394
Qy 390 DAVLKELDEYDVCPTFLGGDMSSALNPTGIYFELPNNACDIYMAVQSFALDVN 449
Db 395 AAGANML-----PGANGVWSPMFPDHTIRIVYANNHQPPTTYVEDAAIPGSKW 447
Qy 450 TSAATKLAPEGNNGRIDAIDISTGRTLSAERPAANYSPVLTAGVFNNGGTDYFR 509
Db 448 LGSAFKTILASQOMKLSAVVVDYGVAMDYKTEQPLIGGLVLTAGGLVFNNGEGNGLFRA 507
Qy 510 LSOETETLMQARLATVATGQALSYELDVOYIAIGAGLTYGTQOLNA 557
Db 508 FDSATGKKLWEPQCGAGVNAVSVWYVWGXOYIAVAAAG-QYARLOA 554

RESULT 15

US-10-369-493-15969
Sequence 15969, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 15969
LENGTH: 592
TYPE: PRT
ORGANISM: Xanthomonas campestris
US-10-369-493-15969

Query Match 15.7%; Score 482.5; DB 14; Length 592;
Best Local Similarity 26.4%; Pred. No. 1.1e-33;
Matches 165; Conservative 105; Mismatches 227; Indels 129; Gaps 28;

Qy 34 ANPPAGEWINGRNOENYHSPLOITADNVGQOLVW-----ARGEMAGAVQVTPMIH 87
Db 14 SNP--DNWGVGRFPALTRHSPFLAEINRDVYKULKMEKMTDTRHEG-----QGLVI 66
Qy 88 DGVWYL--ANDGVYQIALDAQT---GDLIWEHRQLP---AVA---TNAQGRRGV 134
Db 67 GSIMYMAVAYPNV-VALDLASQDGGKVLWKYPPQODERSVAVACCDTVN-----RGA 119
Qy 135 ALYGTSLYFSSWMDHLIALDMETGOVFPVERSGEGDGLTSNTGPIVANGVIAG--STC 193
Db 120 SYADGKLVFGSLGCVIALDIAKTSKEVWKQKLGHPDGEIT-TVAPIIADGKVIYAGISGN 178
Qy 194 QYSPYGCIFSGHDSATGEELWNR-----NHPIPOGEEG--ETWGNDF 234
Db 179 EFGVLG-RVAAYVNLADGQAWSCDAAGKDSICLGAUFNNKAPDHQGLDGLGVTFPND- 236
Qy 235 EARMWTG---VWGQITTYDPTVNLVFGSTGVGPASETORCGTGGTLVGTNTRFAY 286

Db 237 --EWKRGGAAMGWYSYDPLKLYGTGNPGLMSPSYR--CGKTSHEECNNGEHDKMS 292
 QY 281 NTRFAVRPDNGEIVYMRHOTLPRDNWDOECTFEMWVANVDVQPSAEMEGLRAINPNATGE 340
 Db 293 MTLFARKIDTGEAVWGYSQKTFPDQMDYDGINELVDLTI-----DGKEV--PSVVQFD 344
 QY 341 RRVLTGAPCKTGTWMSFPAASGEFLMARDNTYNTMTASID-ETGL-VTVNEDAVLKELDV 398
 Db 345 R-----NGFAYVLDRRDGTLLRAHKFVPANWAEIRIDMKTGRPYKVAHSPL-ERGK 394
 QY 399 EYDVCPTFLGGRDSSAALNP-DTGIYFLPLNACYDIMAVDQFSALDYYNTSATAKLA 457
 Db 395 KVQAFPSAMGGKQOQPCSVDPANSVAFFCGTNNHMKELPQERGNITWGLPYVFANVMKK 454
 QY 458 PGFEN-MGRIDALDIDSTGRTLWSAERPANYSPVLTAGGVFNFGTDRYFRALSOBTGE 516
 Db 455 PNERGALGIYKAPFVVGSKKWEIKEFPVWSGTLVTDGLVFGTLDGWFRAVDKDTGK 514
 QY 517 TLWGRLATVATGCAISYELDGVQYIAI-----GAGGLTY 551
 Db 515 KLMEMKLPSGILGNPIAYKANGHQYAVFSGIGWIGLPVAAGLDPADPYGALGAAGLAF 574
 QY 552 GTQNLAPLAEAIDSTSVGNAIYVFAI 577
 Db 575 GA-----GFDKIPILGKWHTFRI 592

Search completed: November 18, 2004, 05:23:37
 Job time : 147 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 18, 2004, 04:29:32 ; Search time 197 Seconds
(without alignments)
1691.076 Million cell updates/sec

Title: US-10-802-682-8
Perfect score: 3069
Sequence: 1 MNPITLRTSAVALTLTAPA.....AFAIDSTSVGNALYVFPALPQ 579

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1673	54.5	608	2 Q93RE9	Q93RE9 pseudogluc
2	680.5	22.2	575	2 Q89GX5	Q89GX5 bradyrhizob
3	629	20.5	554	2 Q89GY8	Q89GY8 bradyrhizob
4	624.5	20.3	557	2 Q6SEZ4	Q6SEZ4 uncultured
5	624.5	20.3	557	2 AAR38428	AAR38428 uncultured
6	624.5	20.3	724	2 Q6N4Z5	Q6N4Z5 rhodospseudo
7	624.5	20.3	724	2 CAE28629	CAE28629 rhodospseu
8	623	20.3	698	2 Q9KH03	Q9KH03 alcaligenes
9	612.5	20.0	690	2 Q8GR64	Q8GR64 pseudomonas
10	597	19.5	708	2 Q8ED COMTE	Q8ED COMTE pseudomonas
11	594	19.4	595	2 Q88JH0	Q88JH0 pseudomonas
12	594	19.4	695	2 Q89G12	Q89G12 pseudomonas
13	593.5	19.3	602	2 Q89G12	Q89G12 pseudomonas
14	582	19.0	588	2 Q6W1W7	Q6W1W7 rhizobium s
15	582	19.0	588	2 AAQ87251	AAQ87251 rhizobium
16	574.5	18.7	757	1 DHET GLUTOX	DHET GLUTOX gluconobact
17	568	18.5	601	2 Q92W79	Q92W79 rhizobium m
18	565	18.4	629	2 Q9AC08	Q9AC08 methylobact
19	558	18.2	601	2 Q9EYW8	Q9EYW8 rhizobium m
20	556	18.1	691	2 Q9AF95	Q9AF95 pseudomonas
21	553.5	18.0	586	2 Q89D06	Q89D06 bradyrhizob
22	553	18.0	739	1 DHET ACEBU	DHET ACEBU pseudomonas
23	552	18.0	623	1 EXAA PSBAE	EXAA PSBAE pseudomonas
24	551.5	18.0	738	1 DHET ACEPO	DHET ACEPO pseudomonas
25	548.5	17.9	626	1 DHM1 METOR	DHM1 METOR methylobact
26	543	17.7	582	2 Q89X11	Q89X11 bradyrhizob
27	542.5	17.7	622	2 Q75G65	Q75G65 pseudomonas
28	542.5	17.7	622	2 BAD1158	BAD1158 pseudomonas
29	542	17.7	750	2 Q6SFC4	Q6SFC4 uncultured
30	542	17.7	750	2 AAR38298	AAR38298 uncultured
31	541.5	17.6	631	2 Q88JH5	Q88JH5 pseudomonas

32	538	17.5	601	2 P71509	P71509 methylobact
33	537	17.5	623	2 Q9AGW3	Q9AGW3 pseudomonas
34	535.5	17.4	626	1 DHM1 METEX	DHM1 METEX methylobact
35	533.5	17.4	631	2 Q8VPS5	Q8VPS5 pseudomonas
36	530	17.3	750	2 Q8RTS8	Q8RTS8 uncultured
37	511.5	16.7	742	2 Q53362	Q53362 acetobacter
38	504.5	16.4	633	2 Q24759	Q24759 hyphomicrob
39	498	16.2	631	1 DHM1 PARDE	DHM1 PARDE paracoccus
40	493.5	16.1	661	2 Q8PPN7	Q8PPN7 xanthomonas
41	491.5	16.0	742	1 DHET ACEAC	DHET ACEAC acetobacter
42	491.5	16.0	742	2 Q8KZS8	Q8KZS8 acetobacter
43	488.5	15.9	600	1 XOXF PARDE	XOXF PARDE paracoccus
44	471	15.3	695	2 Q834G0	Q834G0 pseudomonas
45	468.5	15.3	599	2 Q8RMH5	Q8RMH5 methylobact

ALIGNMENTS

RESULT 1	ID	Q93RE9	PRELIMINARY;	PRT;	608 AA.
AC	Q93RE9;				
DT	01-DEC-2001 (TREMblrel. 19, Created)				
DT	01-DEC-2001 (TREMblrel. 19, Last sequence update)				
DT	01-OCT-2003 (TREMblrel. 25, Last annotation update)				
DE	Alcohol dehydrogenase.				
GN	Name=adh;				
OS	Pseudoglucobacter saccharotogenes.				
OC	Bacteria; Pseudoglucobacter.				
OX	NCBI_Taxid=133921;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=IFO 14464;				
RA	Shibata T., Saito Y.;				
RL	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AB046580; BAB62258.1; .				
DR	GO; GO:0006118; P:electron transport; IEA.				
DR	InterPro; IPR002372; POQ repeat.				
DR	InterPro; IPR01047; Quin_alc_DH_like.				
DR	Pfam; PF01011; POQ; 2.				
DR	SMART; SM00564; POQ; 5.				
SQ	SEQUENCE 608 AA; 65101 MW; 0ACEG97AE11BA570 CRC64;				
Query Match	54.5%; Score 1673; DB 2; Length 608;				
Best Local Similarity	52.5%; Pred. No. 1,3e+106;				
Matches	315; Conservative 94; Mismatches 167; Indels 24; Gaps 6;				
QY	2 NPTTLRTSAVALTLTAPAPRQ-----VTPITDELIANPPAGEWI 42				
DB	9 NVAGLALSTALIASLSCP-AFAQHDANAAAPSKAGQAIENFPVTADDLAKKUPANWP 67				
QY	43 NVGRNOEYRSPPLQITADNVGOLVYARGMEAGVQVTPMTHDGYMLANPGDVIOA 102				
DB	68 IIRGVYQGWGSPDLQINKXNVGDLQVWSTMEGSGEALAYNGVIFIGNTVDVIOA 127				
QY	103 IDAQTGDIWEHRQLPAVAT-LNAQDQKRGVALYGTSLYFSSNDNLIALDMETGVV 161				
DB	128 IDGKTGSLIMEYRRLPSAKFINSIGAKRSIALFGDKYFVSGVNDNFVALDKTGTLA 187				
QY	162 PVERSGSGEDLSTNTGPIYANGIVAGSTCOVSPVCFTSGHDSAGELMNHFTPO 221				
DB	188 WETNRGGQVEBVGVAISGPVIVDGVILGSGTCQSGFCVYTGDAESGEELMRNTPFR 247				
QY	222 PEEBDEDTWGN-DEPARMTGVMGQITYPVNTLVFYSGTGVGPASETORSTPGTLYGT 280				
DB	248 PEEBDDTGWGAPYNNRNMTGAMGQITYPDLIVYSGTAGPASEVQRTGEGTLAGT 307				
QY	281 NTRFAVVDPTGELIVRQOTLPRDNDQDCTPEMMVAVNDVPSAMEELRLINPAAATGE 340				
DB	308 NTRFAVVKRTGEVVRKQTLPRDNDQDCTPEMMVAVNSVNPDAKADQMSVGVANVRGE 367				
QY	341 -RRVLTGAPCKTGTWMSFDAASGEFLMARDTNNTMISIDETGLIVNEDAVLKELDVE 399				

```

Db      368 TRKVLGVPECKTAVQAFDAKTDGYFWSKATVEQNSIASIDDTGLVTVEMDILKEPGKT 427
Qy      400 YDVCPTFLGGRDSSALNPDTGTIYFLPLNACYDIMAADOESALDVNTSATAIADAG 459
Db      428 YNCCPTFLGGRDMPKAGYLPKSNLYIPLSNACYDVMARTEKTRPADVYNTDAILVLANG 487
Qy      460 FENMERIDAIDISTGRTLMSAERPAANYSFVLSTAGGVFNQGTDRYPALSOETGETILM 519
Db      488 KTMNGRVDAIDLATGETKWSYETRALYDPVLTGTGDLVFEVGGIDRDFRALAESGKEW 547
Qy      520 QARLATVATGQATSYELDGVQYIAGAGLTYGTQANAPLAAIDSTSVGNALYVAPALPQ 579
Db      548 STRLPAGVSGITTSYIDGNQYAVVSGG-SLGGFTPGFTTPVDASAGANGIYVAPALPE 606

```

RESULT 2

```

Q89GX5 PRELIMINARY; PRT; 575 AA.
ID 089GX5
AC 089GX5
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE B116220 protein.
GN OrderedLocustNames=b116220;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiyama T.,
RA Sasamoto S., Watanabe A., Idegawa K., Iriyuchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpou S., Tsurunaka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005957; BAC51485.1; -.
DR HSSP; Q924J7; 1FUG.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR002372; PQQ_repeat.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR Pfam; PF01011; PQQ; 3.
DR SMART; SM00564; PQQ; 6.
KW Complete proteome.
SQ SEQUENCE 575 AA; 63421 MW; F428035287AD54D CRC64;

```

Query Match 22.2%; Score 680.5; DB 2; Length 575;
 Best Local Similarity 32.6%; Pred. No. 2,5e-38;
 Matches 197; Conservative 99; Mismatches 226; Indels 83; Gaps 27;

```

Qy      12 AVLLLTAPAF-----QVPTIDELLANPAGKINVGNOENYHSPILQITADNVGQL 67
Db      13 AFTCLASTAAGAGPIENYSPTAQRLENPEPSNMWYRRTYDQGSPLDQINTSVKGL 72
Qy      68 QLVW--ARGWAGAVQVTPMIDGVWYLANPGDIVIALDAQTGDLIWEHRRLPAVATLN 125
Db      73 TPVWTFATGVBEH-EAPFIVNNGVMEVATPMGVQVALANAKTDEYWRKRLP--DDL 129
Qy      126 AQGDRKRGVALYGTSLYFSWMDHLALDMETGOVYFVDERGSGEGSLNTGPIVANG 185
Db      130 QLHPTSGVGLMWDKLTATTDHVAVALDAKTKGVWVDKRVQDYKKG-QYMTLMLIVDG 188
Qy      186 -VIVAGSTCYSPGCFISGDSATGELWRNHFIPQGEDETV-GNDFE-----ARWM 239
Db      189 KVIIVGSGGFGFRG-VVAIVDAKDKELMRTYTIIGEGEPGHDITWGDDMKNGGSSAM 247
Qy      240 TGVWQITVDPVNLVFFYSGTGVGP-ASEFGRTPGGTLVGNTRAVRPDGEIYWRQ 298
Db      248 TG-----NYDKDKTIYGVGNAPFPGETH---PGDNLY-TSSVLADLPNNKIKTYHQ 298

```

```

Qy      299 TLPRDNDQECTFERMVAVNDVQPSAMEGLRAINPAATGERR---VLTGAPCKTGM- 354
Db      299 YHONDSMDWBEVEZAPML--IDLQRDG-----RSIKSLVHGRDAIFWVLERTPTKINYVA 351
Qy      355 -WSPDASGEFLMARDNTNYNMISID-ETGLVTVNEDAVLKELDVEYDVCPTFLGGRDW 412
Db      352 GMPFVS-----IDVWKGIDAETGKPIV-DPAKPIVIGKREVCPSLMGCKDM 397
Qy      413 SSAALNPDTGTIYFLPLN-NNC---YDIMAADOESALDVNTSATKALAPGEENGRID 467
Db      398 PSAVYSOKTGLIVYPAVENFCGFTGKVALKGEELWLTGKPEDIGLTKIPGADHFELO 457
Qy      468 AIDISTGRTLMSAERPAAN-YSPVLTSTAGGVFNQGT-DRYPALSOETGETIMOARLAT 525
Db      458 AMDPVYTGKWKQHNFPKSGQLFGSVTATAGDLIFRAGINDNFRPAKTELLMECKTNS 517
Qy      526 VATGOAISYELDGVQYIAGAG-GL-----TYGTQANAPLAAIDSTSVGNAL 572
Db      518 GIMGMFVSYEIDGTQYIAGSGWGVDAQRIDQALVTNIGIEANVP-----QGGVI 568
Qy      573 YVPEAL 577
Db      569 WVFAL 573

```

RESULT 3

```

Q89GY8 PRELIMINARY; PRT; 554 AA.
ID 089GY8
AC 089GY8
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Exa protein.
GN Name=exa; OrderedLocustNames=b16207;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiyama T.,
RA Sasamoto S., Watanabe A., Idegawa K., Iriyuchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpou S., Tsurunaka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005957; BAC51472.1; -.
DR HSSP; Q924J7; 1FUG.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR002372; PQQ_repeat.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR Pfam; PF01011; PQQ; 2.
DR SMART; SM00564; PQQ; 6.
KW Complete proteome.
SQ SEQUENCE 554 AA; 60554 MW; CB8B797BAD9B061 CRC64;

```

Query Match 20.5%; Score 629; DB 2; Length 554;
 Best Local Similarity 30.5%; Pred. No. 8.4e-35;
 Matches 184; Conservative 101; Mismatches 229; Indels 90; Gaps 23;

```

Qy      10 SAALLLTAPAFQVPTIDELLAN-PKAGKINVGKNOENTRHSPLTQITADNVQQL 68
Db      7 AASLVLAATYANQOT--TEQLVKGATDTSNVNTGMGYNLRFTSLNINKDTVNLV 63
Qy      69 LVWARGM-EGAGVQVTPMIDGVWYLANPGDIVIALDAQTGDLIWEHRRLP-VAITNA 126
Db      64 PWNVYSPRDDRSESSQLYVGVYIYYS-HNATPAVDAKXGKQKSKIKIYPAETPEIYC 122
Qy      127 QGDRKRGVALYGTSLYFSWMDHLALDMETGOVYFVDERGSGEGDLTSTTGPVANGV 186

```

```

Db      123  CGIIRBGAALYDGVFFETITDANVIALDADGKEIMWQKADKEG-SMTTAPVADGV 181
Qy      187  IVAG-STCQSPYGCIFSGHDSATGEELWENHPIPOGEEGDETWGNDP-----EARMWT 240
Db      182  VIRGISCAEFGTRG-FIDGWDPAATGKILMRTHSIPSPDEPGDWTMGDTWKGCSSTWIT 240
Qy      241  GWVGQITVDPTLVFVSTGVGPASBETRGTPGGTLVGTNRFAVRPDIGEILWHSQTL 300
Db      241  G-----SYDELNTVTVGIGNPGPFNSAVR-PGDNLV-TCSVLADPPTGKIKMHYQPS 292
Qy      301  PRDNMOECTFEMWVA--NVDVOPSAMEGLRAINPNAATGERVLTGAPCKTGTWMSFD 358
Db      293  PNPFPDYSAEVAWVLADMNVECKPT-----KVLMDAN-RNGFFVYLD 333
Qy      359  AASGEFLMADPTVNTNIAISID-ETGLVTVNEDAVLKEIDVEYD-----YCFPLGG 409
Db      334  RTNGKLLAANPYKVNATGVMKTRPT-----ETDVSXDAEKKVTVYPSILGG 385
Qy      410  RDWSSAALNPDGTIFLPLNNACYDIWVDOEFSALDVNTSATKLADGFENM----- 463
Db      386  KNMEPMSFNPOTGL-----AVANTLAFGKTKABPV-----TFQGEWYLGMDLTPM 433
Qy      464  -----GRIDAIDISTGRTLWGAERPAANYSPVLSTAGGVVNGGTDKRYFRALSOETG 516
Db      434  EFGDGARGHKLKALDPMTGAKWEPADIPRFSGLSTAGGVVFTGALTGFEAFADATGK 493
Qy      517  TLWQARLATVATGQALSYELDGVQYTAIGAGLTYGTQNALPILA-BAIDSTVSQNALIYF 575
Db      494  KMGQFQSGSIBEQPVTWQDDVQYIAVTSG--YGGVTSLSFGDERLAKVPFGSLWVF 550
Qy      576  ALPQ 579
Db      551  AVKQ 554

```

RESULT 4

```

ID 06SE24 PRELIMINARY; PRT; 557 AA.
AC 06SE24;
DT 05-JUL-2004 (TEMBLrel. 27, Created)
DT 05-JUL-2004 (TEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TEMBLrel. 27, Last annotation update)
DE PQQ enzyme repeat domain protein.
GN ORFNames=EBAC080-L028H02.93;
OS uncultured bacterium 582.
OC Bacteria; environmental samples.
OX NCBI_Taxid=257402;
RN [1]
RP SEQUENCE FROM N.A.
RA DeLong E.F.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Heidelberg J.F., Eisen J.A., Nelson W.C., DeLong E.F.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY458649; AAR38428.1;
DR InterPro: IPR002372; PQQ_repeat.
DR InterPro: IPR011047; Qun_alc_DH_like.
DR Pfam: PF01011; PQQ; 2.
DR SMART: SM00564; PQQ; 7.
SQ SEQUENCE 557 AA; 61170 MW; 10A36DDEF2D15A08 CRC64;

```

Query Match 20.3%; Score 624.5; DB 2; Length 557;

Best Local Similarity 32.3%; Pred. No. 1.7e-34; Matches 189; Conservative 82; Mismatches 250; Indels 65; Gaps 22;

```

Qy      12  AVLLTAPAAFAQVPTTDELLANPAGEWINGRQENYRHSPLQTITADNVGQLQVW 71
Db      15  ALSLIAPETTLAQT--SDLINDTKDTTTLITGWMYGGQTRHSGLAGINDNIGRLRPW 72
Qy      72  ARGMEAGAVVT-PMHIDGVWYLANPGDVYQALDAQOTGLIY---EHRRLPVAVALTA 126

```

```

Db      73  TYSLSDTRGQETFLPHIDGKMYVTTHASTM-ALDLVTRGQIKWTLVDVDPADTPRVACC-- 129
Qy      127  QCDRRRGVALLYGTSLYFSSWMDNLIALDMETQGVVDFVERSGEDGLTSNTTGPVANGV 186
Db      130  -GIVNRGVALYEGKFFRTLLDAHVIALDAETGNELMRQSLDKHTGY-SMTVAPWMAADGV 187
Qy      187  IVAG-STCQSPYGCIFSGHDSATGEELWENHPIPOGEEGDETWGNDP-----EARMWT 240
Db      188  LITGISCEYGRG-YLEGIDPDIGVRVWRVYTTIPAPGSPSEITWQDGDAMTNGAPTW 246
Qy      239  MTGVGQITVDPTLVFVSTGVGPASBETRGTPGGTLVGTNRFAVRPDIGEILWHSQTL 300
Db      247  LTG-----SYDELNTVTVGIGNPGPFNSAVR-PGDNLV-TCSVLADPPTGKIKMHYQPS 292
Qy      239  PRDNMOECTFEMWVA--NVDVOPSAMEGLRAINPNAATGERVLTGAPCKTGTWMSFD 358
Db      229  TLPDNMOECTFEMWVA--NVDVOPSAMEGLRAINPNAATGERVLTGAPCKTGTWMSFD 337
Qy      359  AASGEFLMADPTVNTNIAISID-ETGLVTVNEDAVLKEIDVEYD-----YCFPLGG 409
Db      338  RATGELIAPKRVOKLWADSIDLETTRPKETDVAMKARSGEQLTFMPSALGCKWSPMA 397
Qy      417  LNPDTGITYFLPLNNAICYDIWVDOEFSALDVNTSATKLADGFENMGRIDAIDISTGRT 476
Db      398  YNPKTRITLFAVNTKRGKMYKAVEQYRP-GVFFGAEFSMDPDQDGELEAFDPMTGVV 456
Qy      477  LMSAERPAANYSPVLSTAGGVVNGGTDKRYFRALSOETGTLWQARLATVATGQALSYEL 536
Db      457  KMSDSIALPRYSGVLTSGGLVFTGNMDGEFAEATGDKWYVTSQSGIIGQPVWIER 516
Qy      537  DGVQYIAI-GAGS--LTYGTQNALPILAEAIDSTVSQNALIYF 577
Db      517  BGQYITVANGSGGVYSLFSGD-----ERLASVPAAGNIWTFSI 555

```

RESULT 5

```

ID AAR38428 PRELIMINARY; PRT; 557 AA.
AC AAR38428;
DT 02-MAR-2004 (TEMBLrel. 27, Created)
DT 02-MAR-2004 (TEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TEMBLrel. 27, Last annotation update)
DE PQQ enzyme repeat domain protein.
GN EBAC080-L028H02.93.
OS uncultured bacterium 582.
OC Bacteria; environmental samples.
OX NCBI_Taxid=257402;
RN [1]
RP SEQUENCE FROM N.A.
RA DeLong E.F.;
RL "Monterey Bay Coastal Ocean Microbial Observatory environmental clone sequencing."
RT Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Heidelberg J.F., Eisen J.A., Nelson W.C., DeLong E.F.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY458649; AAR38428.1;
DR EMBL: AY458649; AAR38428.1;
SQ SEQUENCE 557 AA; 61170 MW; 10A36DDEF2D15A08 CRC64;

```

Query Match 20.3%; Score 624.5; DB 2; Length 557;

Best Local Similarity 32.3%; Pred. No. 1.7e-34; Matches 189; Conservative 82; Mismatches 250; Indels 65; Gaps 22;

```

Qy      12  AVLLTAPAAFAQVPTTDELLANPAGEWINGRQENYRHSPLQTITADNVGQLQVW 71
Db      15  ALSLIAPETTLAQT--SDLINDTKDTTTLITGWMYGGQTRHSGLAGINDNIGRLRPW 72
Qy      72  ARGMEAGAVVT-PMHIDGVWYLANPGDVYQALDAQOTGLIY---EHRRLPVAVALTA 126
Db      73  TYSLSDTRGQETFLPHIDGKMYVTTHASTM-ALDLVTRGQIKWTLVDVDPADTPRVACC-- 129
Qy      127  QCDRRRGVALLYGTSLYFSSWMDNLIALDMETQGVVDFVERSGEDGLTSNTTGPVANGV 186

```

```

Db      130  -GTVNGVAVLYEKKVRRITLDAIVILADLDETGEMLRTGSDIKHKGY-SMTVAAPMAAGCV 187
      187  IVAG-STCOYSPYCGTISGHDSATGEELMRNHFIPQGEGBETW--GNDPEAR-----W 238
      188  LITGISGEGYGLRG-VLEGYDPTGVVRWRYTIPAPGEGSEITWQDGDADWTGCAPTW 246
      239  MTCWVGQITVDPVTNIVFGSGTGVGPASETQGTGCTGYLTNTRPAVDPDGEIWMRQ 298
      247  LIT-----SYDEPLNTVFWG-TG-NPSSNMAATRPEDNLYASSI-LALDPKGTIKHKHQ 298
      299  TLEPRDWDDECTEMVAVANVQPSAEMEGRAINENNAATGERRVLTGAPCKTGTMMSFD 358
      299  TTPNDPEFDEPTVELVHADID-----GRKVLMAQAN-ANGGEFFVYLD 337
      359  AASGEELMA-RDNTYNNMAISID-ETGLVTNVEDAVLKELDVEDYDCPEFLGGRMSSAA 416
      338  RAIGELIAAKFVYDKLNMADSIDLETRRPKETDVAAKASGRIITFWPALSGLKMSFMA 397
      417  LNPDTGIYFLPLNNACYDIMAVIDEFSALDVNTSATAKLAPGFNMGRIDAIDISTGRT 476
      398  YNKKTITFANTMKFGVHYKAVEPQYRP-GVFFYFGAEFSMDWPDGDRGELRAFDPMTGIV 456
      477  LMSAEPPAANYSPVLTAGGVVFNAGTDTRYFRALSGETGETILMQARLATVAAGQALSYEL 536
      457  KMSDSIAIPRYSGVLTGGGLVFTGNMDEFEAFNAETGDKWVTYSTGSGIIGQPYTWER 516
      537  DGVOYIAI--GAGG--LTYGTQLNAPLAELIDSTSVGNAIYFAL 577
      517  EGGQYITVANGSGGVSLFSGD-----ERLASVAGGNIMTFST 555

```

RESULT 6

```

ID      06N425      PRELIMINARY;      PRT;      724 AA.
AC      06N425;
DT      05-JUL-2004 (TREMblrel. 27, Created)
DT      05-JUL-2004 (TREMblrel. 27, Last sequence update)
DT      05-JUL-2004 (TREMblrel. 27, Last annotation update)
DE      Alcohol dehydrogenase precursor (EC 1.-.-.-).
GN      OrderedLocustNames=RP31188;
OS      Rhodospseudomonas palustris.
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC      Bradyrhizobiaceae; Rhodospseudomonas.
OX      NCBI_TaxID=1076;

```

```

RN      SEQUENCE FROM N.A.
RP      STRAIN=CGA009 / ATCC BAA-98;
RX      PubMed=14704707; DOI=10.1038/nbt923;
RA      Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA      Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA      Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA      Harrison F.H., Gibson J., Harwood C.S.;
RT      "Complete genome sequence of the metabolically versatile
RT      photosynthetic bacterium Rhodospseudomonas palustris."
RL      Nat. Biotechnol. 22:55-61(2004).
EMBL: BX572603; CAE28629.1; -.
CO      GO:0016491; F:oxidoreductase activity; IEA.
DR      InterPro: IPR000345; CytC_heme_BS.
DR      InterPro: IPR009056; Cytochrome_C.
DR      InterPro: IPR003088; Cyt_C1.
DR      InterPro: IPR002372; PQQ_repeat.
DR      InterPro: IPR011047; Quin_alc_DH_like.
DR      Pfam: PF00034; Cytochrom_C_1.
DR      SMART: SM00564; PQQ_5.
DR      SMART: SM00564; PQQ_5.
DR      PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
KM      Complete proteome; Oxidoreductase; Signal.
FT      SIGNAL 1 45 Potential.
SQ      SEQUENCE 724 AA; 78144 MW; 474551BA6FCBCE98 CRC64;

```

Query Match 20.3%; Score 624.5; DB 2; Length 724;
Best Local Similarity 29.5%; Pred. No. 2; Se-34;

Matches 184; Conservative 100; Mismatches 249; Indels 91; Gaps 24;

```

CY      8  RTSAAYL---LTPAPAPAO-----VTPITDE--LLAN-PPAGEWINGRNOEN 50
      26  RLGAALMAALLTPVAVAADSDSGSAAHITAVGAIVASAIIVANVTXTDMSPYGLDYAE 85
      51  YRHSPLTQTADNVGQLQVW--ARGMEAGAVOTPMIHGVNTYLANPDGVYQALDAQTG 108
      86  TRFSKLQITENVKQGLQMSYSLGSENG-VETAPVVDGIMVYVVASVVAIDITRTG 144
      109  DLIVHRKQLPAVATLNAQDR-----KRGVALYGSILFSSMDNHLALDDETQGVVF 162
      145  KKLMTFDPKVD---HSKGYRGCCDVNRGVNLYKSKYVGAIDGRLLDLDAATSKWM 199
      163  DVERSGEDGLTSNTTGPPIVANGVIYAGS-TCQYSPYCGTISGHDSATGEELMRNHFIP- 220
      200  EIDFLIDHESYITLTAQVAVNGKVIYVINGGAEYGARQ-VYTAVIDETGKQAWRMKFTVPG 258
      221  -----QGEEDDEFWGNDFEAKMMTG---VMQITVDPVTNIVFGSGTGVGPASET 268
      259  DPKRPFDESNEMKAKTM-DPAKRWLWGGGGTAMDITTFDPLNLIVGTGNGSEPNRH 317
      269  QGTGPGGLTGYTRFAVRPDTGSIYVRHQTLPRDWDDECTEPMNVAN--VDVQPSAEM 326
      318  LRSFAGGDNLVLAISVALNADTGKVVWHYGETPGDMDVYSTQPMILADIAIDGKP---- 373
      327  EGLRPAINPAATGERRVLTGAPCKTGTMSPFDAASGEFLMARDTNTNKLASIDETGL-V 385
      374  -----RKYLILAP-KNGFFVIDRTDGFISAKNFVDVNMATGYDSNGRPI 418
      386  TVNEDAVLKELDVEDYDCPEFLGGRDWSAALNPDTGIYFLPLNNACYDI--MAVDQ-- 440
      419  EVPE--ARSDKSFDAIPGPIGAHWHMPSFNPQGLVYLPAGQVPMVLTGSKALTONK 475
      441  ---EFSALDVNTSATAKLAPGFNM-GRIDAIDISTGRTLMSAEPPAANYSPVLSTRA 494
      476  MEPRFSGTGMVNGFVLNAVPP-KNLPFGRILAMPVQCKEYWRAYVSPMNGGTLTTA 534
      495  GGVVFNAGTDTRYFRALSGETGETILMQARLATVAAGQALSYELDGVOYIAI-GAG-GTYGT 553
      535  GNLVFOGTADRFAVNAKKEKIMESPGLTGAAVAPATYVMDGVQYVSLAVMGVYFGI 594
      554  QLNAPLAELIDSTSVGNAIYFAL 577
      595  SARATEEA-----PGVYTFPAV 612

```

RESULT 7

```

ID      CAE28629      PRELIMINARY;      PRT;      724 AA.
AC      CAE28629;
DT      02-MAR-2004 (TREMblrel. 27, Created)
DT      02-MAR-2004 (TREMblrel. 27, Last sequence update)
DT      02-MAR-2004 (TREMblrel. 27, Last annotation update)
DE      Alcohol dehydrogenase precursor (EC 1.-.-.-).
GN      RP31188.
OS      Rhodospseudomonas palustris.
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC      Bradyrhizobiaceae; Rhodospseudomonas.
OX      NCBI_TaxID=1076;

```

```

RN      SEQUENCE FROM N.A.
RP      STRAIN=CGA009 / ATCC BAA-98;
RX      PubMed=14704707;
RA      Larimer F.W., Chain P., Hauser L., Lamerdin J., Malfatti S., Do L.,
RA      Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA      Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA      Harrison F.H., Gibson J., Harwood C.S.;
RT      "Complete genome sequence of the metabolically versatile
RT      photosynthetic bacterium Rhodospseudomonas palustris."
RL      Nat. Biotechnol. 22:55-61(2004).
EMBL: BX572603; CAE28629.1; -.
KM      Oxidoreductase; Signal.

```

FT SIGNAL 1 45 Potential.
SQ SEQUENCE 724 AA; 78144 MW; 474551BA6FBC98 CRC64;
Query Match 20.3%; Score 624.5; DB 2; Length 724;
Best Local Similarity 29.5%; Pred. No. 2, 3e-34;
Matches 184; Conservative 100; Mismatches 249; Indels 91; Gaps 24;
QY 8 RTSAAYLV---LTPAPAAFAO-----VTPIDE--LLAN--PPAGWINGYGNQEN 50
DB RLGAALMALMLTPAAALADSKGSAHRAVGAADSAIVANVTTCMDPSYGLDYAE 85
QY 51 YRHSPLTQITADNVGQQLVW--ARGENAGAVOTPMIHGVVYLANPGDVIQALDAQTG 108
DB TRFSKIDQITENTENKQGLQMSYSLASERG--VEATPVVDGIMYVYASVVAHIDTRIG 144
QY 109 DLWEHRROLPAVATLNAOGDR-----KRGVALYGTSLYFSSWNHLLALDMEGTQVVF 162
DB 145 KXIMTDPKVD-----HSKGYGCCDVYVNGVALYKGYFVGAYDRLIALDATSCKAM 199
QY 163 DVERSGSDGLTSNTGPIVANGVIYAGS--TCQYSPYGPISGHSATGSELMRNHFIP- 220
DB 200 EIDTLIDHEHSYITITGAPRVENGKVVITGNAGAEYARG--YVAYDAETGKQARWFTVPG 258
QY 221 -----QGEEDGEMGNDFEARWMTG---VMQIITDPVTNLVPGYSGTGVGASSET 268
DB 259 DPKRPEDESEMEKAATW--DPAGKMWLNGGGGTAMDTITFDPLNLIVGTGNGSEFWNH 317
QY 269 QRGTPGTYTGTNTRFAVPTGEIIVWRHOTLPRDNMDQCEPEMNVAN--VDVQPSAM 326
DB 318 LRBPAGGDNLYLASIVANLADTKYVWHYGETPGDMDVTSOPMLAIDAKR----- 373
QY 327 EGLRNPNAATERRVLTGAPCKTGTMSFDAASGEFLMARDTNTNMLASIDETGL-V 385
DB 374 -----RKVILHAP--KNGFFVIDRTDKFISAKNFVDVNNATGYSNGRPI 418
QY 386 TVNEDAVLKEIDVEYDVCFTFLGRDMSAALNPDGIYFLPLNNAACYDI--MAVDO-- 440
DB 419 EYVE---ARADSPDAIPGYPGAMHWHMPSFNPQGLVYLPAGQVPPVNLGSKALTONK 475
QY 441 ---EESALDVNTSATKLAPEGENN--GRIDAIDISTGRTLSAERPANYSPYLSTA 494
DB 476 MEFKFGSTTGMNVGFTLNAPV--KNLPFGRLVAMPVQCKEVRAYEYVSPMNGFTLTA 534
QY 495 GGVVNGGDRYRRAISOETGELMQARLATVATGALSYELDGVQYIAGG--GLTYGT 553
DB 535 GNLVFGQIDAGRVAANAKTGERKLWESPLGTGAVAAPATYMDGVQVISIAGMGVFGI 594
QY 554 QLNAPLAEAIDSTVGNAIYFAL 577
DB 595 SARATEREA-----PGTVYTFVAV 612
RESULT 8
Q9KH03 PRELIMINARY; PRT; 698 AA.
ID Q9KH03
AC Q9KH03; PRELIMINARY; PRT; 698 AA.
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DS Tetrahydrofurfuryl alcohol dehydrogenase.
OS Alcaigenes eutrophus (Ralstonia eutropha).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Wauteria.
OC NCBI_TaxId=510;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21125557; PubMed=11222593;
RA Zarni G., Schrader T., Andreesen J.R.;
RT "catalytic and molecular properties of the quinoxinoprotein
tetrahydrofurfuryl alcohol dehydrogenase from Ralstonia eutropha
strain Bo.";
RT J. Bacteriol. 183:1954-1960(2001).
DR EMBL; AF277373; AAF86335.1; -.

DR HSP: Q46444; IKB0.
DR GO; GO:0005746; C:mitchondrial electron transport chain; IEA.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0006118; F:electron transport; IEA.
DR InterPro; IPR000345; CyC_heme BS.
DR InterPro; IPR009056; Cytochrome_c.
DR InterPro; IPR003088; CyC_C1.
DR InterPro; IPR002372; PQQ_repeat.
DR InterPro; IPR01047; Quin_alc_DH_like.
DR Pfam; PF00034; Cytochrome_C; 1.
DR Pfam; PF01011; PQQ; 4.
DR SMART; SM00564; PQQ; 5.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN 1.
SQ SEQUENCE 698 AA; 75471 MW; 9C87F180BD12FB2 CRC64;
Query Match 20.3%; Score 623; DB 2; Length 698;
Best Local Similarity 29.8%; Pred. No. 3e-34;
Matches 182; Conservative 99; Mismatches 249; Indels 80; Gaps 24;
QY 10 SAAVLLTPAPAF-AQVTPITD--ELLANPPAG--EWINGRNOENYHSPLOTITADNV 64
DB 14 AAASVALPAPAFGANAARAVDGAIRAN--EAGTPNPSYGLDYAETRFSLBQVNAAGNV 72
QY 65 GOLQVWARGMEA--GAVVTPMIDGVVYLANPGDVIQALDAQTGDLIWEHRROLPAVAT 123
DB 73 RNLGLAWSYDLESIRGEATPLVVDGVVYASVVAHIDARTGKRLMTYDQVPRDQA 132
QY 124 LNAQGD--KRGVALYGTSLYFSSWNHLLALDMEGTQVFPDVERSGSDGLTSNTGPIV 182
DB 133 YKGCDDVYVRGVALYQGVNPGAFDGRVVALDAATGKRWQDPIVDRSKSYITITGAPRV 192
QY 183 ANG-VIYAGSTQYSPYGCFTSGHDSATGSELMRNHFIP---QGE-----EGDETWGN 232
DB 193 YNGKYLINGAGAEYGVG--YITAYDAETGKQOMRWYTPGDPAPPFENEMAKAAATW--D 250
QY 233 DFEARWMTG---VMQIITDPVTNLVPGYSGTGVGPASETGRRPGG--TYGNTTRFAVR 287
DB 251 PSKKTWINGGGGYVNTNPAFDELNLKMTIGTGNAPMSRKLRSKGGDNLTAASV--VALN 309
QY 288 PDTGEIIVRHOTLPRDNMDQCEPEMNVAN--VDVQSAEVEGLRANPNAATGERVLT 345
DB 310 PDTGVVWHYGETPGDMDVTSQDILADKIDGP-----RKVIL 351
QY 346 GAPCKTGTMSFDAASGEFLMARDTNTNMLASIDETG--LVYNEDAVLKEIDVEYDVC 403
DB 352 HAP--KNGFFVIDRTNGKFIISAKNFVDVNNASGYDKRGPVETPOADTSGPA---DVV 406
QY 404 PTFLGGRDMSAALNPDGIYFLPLNNAACYDINAVDOFSALDVNTSATKLT----- 456
DB 407 GPGFAGHWHMSFHPKLGAFIFAGHYPL--TLADNKEW-----VHNGKDSFEARHGVGN 461
QY 457 -----APGFENGRIDAIDISTGRTLSAERPANYSPVLTAGGVFNGGTDRYFR 508
DB 462 LGMVNAEPKSKEMGRILAMPPLAQKAVWRHDAGCPWNGGTATLTAENLVFOGADRLV 521
QY 509 ALSQETGELMQARLATVATGALSYELDGVQYIAGG--GLTYGTQLNAPLAIDSTS 567
DB 522 AYHATGKLMQAPGSGVVAAPVTYLLDGRQYVSVAVNGGYTG-----LSGRGSDRO 575
QY 568 VGNAIYFAL 577
DB 576 APCRVTYFVL 585
RESULT 9
Q8GR64 PRELIMINARY; PRT; 690 AA.
ID Q8GR64
AC Q8GR64;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DS Quinoxinoprotein alcohol dehydrogenase ADH IIB.
GN Name=qdaa;

OS Pseudomonas putida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;

RA TOYAMA H., FUJII T., AOKI N., MATSUHITA K., ADACHI O.;
RT "Molecular Cloning of Quinolomoprotein Alcohol Dehydrogenase, ADH IIB,
RL from Pseudomonas putida HK5.";
RL Biosci. Biotechnol. Biochem. 67:1397-1400(2003).
DR EMBL; AB091400; BAC15559.1; -.
DR PDB; 1KV9; X-ray; A=23-690.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. .; IEA.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0006118; F:electron transport; IEA.
DR InterPro; IPR001479; Bac_PQQ.
DR InterPro; IPR00345; CytC_heme_BS.
DR InterPro; IPR009056; Cytochrome_c.
DR InterPro; IPR003088; Cyt_CI.
DR InterPro; IPR002372; PQQ_repeat.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR Pfam; PF00334; Cytochrom_G; 1.
DR Pfam; PF01011; PQQ; 3.
DR SMART; SM00564; PQQ; 5.
DR PROSITE; PS00364; BACTERIAL_PQQ_2; 1.
DR PROSITE; PS00130; CYTOCHROME_C; UNKNOWN 1.
SQ SEQUENCE 690 AA; 74969 MW; F5198701D0937613 CRC64;

Query Match 20.0%; Score 612.5; DB 2; Length 690;
Best Local Similarity 29.7%; Pred. No. 1.5e-33;
Matches 181; Conservative 79; Mismatches 237; Indels 113; Gaps 20;

7 LRTSAVLLITAP-AAFAQVPTTDELLANPPEWINGRNQENYRHSPLTQTADNVG 65
5 LKSLMLCLATLALAAVDEAARAEQSGELSLGRVYAEGRFPLQIDASNVR 64
66 QLOLVW-----ARGEKAGAVQVTPIHNGVMTLANPGVIALDAQTDLWERRQLP 119
65 SLGLAWMDLDNTRGDEA-----TPLFHGVITSNWSRVIAVDAASGKELM---RDP 116
120 AVATLNAQGD---RKRGVALYGTSLYFSSMDNHLTALDMENGVQVFDVERSGEDGLTS 175
117 EVAKVKARITSCCAVNRGVALMGDKYVGTLDRLTALDAKTOKALW-----SQ 165
176 NTGP-----IVANGVIVAGSTQYSPYGCFIGSDSANGEEILWNNHFLIPQGE 224
166 QITDPAPFYSITGAPRVKGVKVIIGNGCAEYVVG-FVSAVDADTGKLAIR--FTVPGD 222
225 -----EGDETNGDNFEARNMTG---VMQITVDPTNIVFVGSTGVGPASETQ 269
223 PALPYEHPEIRAKWTQGD--QYKLDGGGGVWDSMAYDPELDLIVGTGSGSPNNREV 280
270 RGTGGGLYGTNTRFAVRPDTEIWRHQTLPNDWDOCTEPMVA--NVDVPSAENE 327
281 RSGGGGNLTLSSILAIRPDTGLAWHYQVTPGSDWDFATQOITLAEINIDGKP----- 335
328 GLAIPNNAATGERRVLITGAPCKTGTMSFDAASGEFLWARDNTYNNMLASIDEGVLTV 387
336 -----RKVLMAQF-KNGFFVYVLDRTNGKLI-----SAKKGKVTW 369
388 NEDAVL-----KELDVEYD-----VCPFLGGRDSSAALNPDITGYFLPLNNAICYDM 436
370 AEFKVDLATGRPAVAPGRVKEKPIVWVSPFGAHNHSFNGGTGLVYIPQEVAGVVR 429
437 AVDOESALDIVNTS---ATAKLAPGEMNGRIDALDISTGRTLSAEPRANYSVLT 493
430 NEGDEFTKRAFNTAGFADATDVPAVVSALLANDPVKQKAMVVPYPTMNGTST 489
494 AGGVVNGGTRFRFALSOETGETLWQARLATVATGSAISYELDGVQYLAIGG-----G 548
490 AGNIVFGTAAAGOMHAYSAKSGALWQFNAOSGIVAAPMTFELAGQYVAIMAGMGVAT 549

OY 549 LTVGTQLNAP 558
DB 550 LTGSGESYNLP 559

RESULT 10
OHED_COMTE STANDARD; PRT; 708 AA.
AC 04644;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Quinolomoprotein ethanol dehydrogenase type I precursor (EC 1.1.99.-)
DE (OH-EH11).
GN Name=chgdh;
OS Comamonas testosteroni (Pseudomonas testosteroni).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Comamonadaceae; Comamonas.
OX NCBI_TaxID=285;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15667;
RX MEDLINE=96184549; PubMed=8654419;
RA Stocrovogel J., Kraayveld D.E., van Sluis C.A., Jongejan J.A.,
De Vries S., Duine J.A.,
RT "Characterization of the gene encoding quinohaemoprotein ethanol
RT dehydrogenase of Comamonas testosteroni.";
RL Eur. J. Biochem. 235:690-698(1996).
[2]
RP SEQUENCE OF 32-54 AND 477-490.
RC STRAIN=ATCC 15667;
RX MEDLINE=95324580; PubMed=7601151;
RA de Jong G.A.H., Geertlot A., Stocrovogel J., Jongejan J.A., De Vries S.,
Duine J.A.;
RT "Quinolomoprotein ethanol dehydrogenase from Comamonas testosteroni.
RT Purification, characterization, and reconstruction of the apoenzyme
RT with pyrroloquinoline quinone analogues.";
RL Eur. J. Biochem. 230:899-905(1996).
[3]
RP CHARACTERIZATION.
RX MEDLINE=86242113; PubMed=3521592;
RA Green B.W., van Kleef M.A., Duine J.A.;
RT "Quinolomoprotein alcohol dehydrogenase apoenzyme from Pseudomonas
RT testosteroni.";
RL Biochem. J. 234:611-615(1986).
[4]
RP CRYSTALLIZATION.
RX MEDLINE=21536088; PubMed=11679760; DOI=10.1107/S0907444901013002;
RA Oubrie A., Huizinga E.G., Rozeboom H.J., Kalk K.H., de Jong G.A.H.,
Duine J.A., Dijkstra B.W.;
RT "Crystallization of quinohaemoprotein alcohol dehydrogenase from
RT Comamonas testosteroni: crystals with unique optical properties.";
RL Acta Crystallogr. D 57:1732-1734(2001).
[5]
RP X-RAY CRYSTALLOGRAPHY (1.44 ANGSTROMS).
RX MEDLINE=21671354; PubMed=1174714; DOI=10.1074/jbc.M109403200;
RA Oubrie A., Rozeboom H.J., Kalk K.H., Huizinga E.G., Dijkstra B.W.;
RT "Crystal structure of quinohaemoprotein alcohol dehydrogenase from
RT Comamonas testosteroni: structural basis for substrate oxidation and
RT electron transfer.";
RL J. Biol. Chem. 277:3727-3732(2002).
[6]
RP FUNCTION: Catalyzes the dye-linked oxidation of primary alcohols
RP to the corresponding aldehydes and the (subsequent) oxidation of
RP the aldehydes to carboxylic acids.
CC -!- COFACTOR: Binds 1 PQQ group, 1 calcium ion, 1 heme and 1 iron ion
CC per subunit.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Periplasmic (Potential).
CC -!- INDUCTION: By ethanol and butanol.
CC -!- SIMILARITY: Belongs to the bacterial PQQ dehydrogenase family.
CC -!- SIMILARITY: Contains 1 cytochrome c domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X81880; CAA57464.1; -.
DR PIR; S62366; S52317.
DR PDB; 1KB0; X-ray; A=32-708.
DR InterPro; IPR002372; Bac_PQQ_repeat.
DR InterPro; IPR000345; Cytc_heme_BS.
DR InterPro; IPR009056; Cytochrome_C.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR Pfam; PF01011; PQQ; 5.
DR SMART; SW00564; PQQ; 5.
DR PROSITE; PS00150; CYTOCHROME_C; 1.
DR 3D-structure; Calcium; Direct protein sequencing; Heme;
KM Oxidoreductase; Periplasmic; PQQ; Signal.
FT CHAIN 1 31
FT SIGNAL 1 31
FT 32 708
FT 619 708
FT METAL 216 216
FT METAL 294 294
FT BINDING 635 635
FT BINDING 638 638
FT METAL 639 639
FT METAL 33 43
FT METAL 44 44
FT METAL 47 52
FT METAL 53 55
FT METAL 56 60
FT METAL 63 64
FT METAL 67 68
FT METAL 72 73
FT METAL 80 82
FT METAL 83 85
FT METAL 86 93
FT METAL 105 107
FT METAL 108 109
FT METAL 110 114
FT METAL 116 118
FT METAL 120 124
FT METAL 125 128
FT METAL 129 134
FT METAL 140 145
FT METAL 145 156
FT METAL 156 160
FT METAL 161 165
FT METAL 167 168
FT METAL 170 175
FT METAL 176 178
FT METAL 181 186
FT METAL 187 190
FT METAL 196 196
FT METAL 202 204
FT METAL 205 206
FT METAL 207 210
FT METAL 213 213
FT METAL 215 217
FT METAL 220 220
FT METAL 222 227
FT METAL 228 230
FT METAL 233 238
FT METAL 244 245
FT METAL 251 257
FT METAL 258 259
FT METAL 262 264
FT METAL 266 269
FT METAL 271 272
FT METAL 279 282

FT TURN 283 286
FT STRAND 287 291
FT STRAND 285 296
FT STRAND 300 303
FT HELIX 305 306
FT TURN 311 314
FT TURN 315 319
FT STRAND 321 323
FT STRAND 326 331
FT TURN 334 335
FT STRAND 347 353
FT TURN 354 355
FT STRAND 356 363
FT TURN 366 367
FT STRAND 375 377
FT TURN 380 385
FT STRAND 392 395
FT TURN 397 398
FT STRAND 401 403
FT STRAND 405 408
FT TURN 410 411
FT STRAND 414 415
FT TURN 420 421
FT STRAND 430 432
FT TURN 433 436
FT STRAND 437 444
FT STRAND 448 451
FT TURN 453 454
FT TURN 457 458
FT TURN 462 463
FT TURN 465 466
FT HELIX 467 469
FT TURN 470 470
FT STRAND 475 477
FT TURN 487 494
FT TURN 495 498
FT STRAND 499 506
FT STRAND 514 517
FT TURN 518 520
FT STRAND 521 525
FT TURN 527 528
FT STRAND 530 535
FT TURN 536 538
FT STRAND 541 546
FT STRAND 556 560
FT TURN 561 562
FT STRAND 563 570
FT HELIX 574 579
FT STRAND 590 596
FT TURN 597 597
FT STRAND 599 599
FT HELIX 621 623
FT HELIX 624 634
FT TURN 635 635
FT HELIX 636 639
FT TURN 642 644
FT HELIX 653 655
FT HELIX 662 662
FT TURN 663 663
FT HELIX 664 668

Query Match 19.5%; Score 597; DB 1; Length 708;
Best Local Similarity 28.9%; Pred. No. 1.5e-32;
Matches 178; Conservative 91; Mismatches 251; Indels 96; Gaps 21;

Oy 16 LTAAPAQVTPIT-----DELLANPP-AGEWIVNGNQENVYRHSPLTQITADN 63
Db 23 LGSAAAPQCTGPAAQAAAVQGVDPFIRANNAARTPDPMTIGVDYMETRYSKLDQIVAA 82
Oy 64 VGQIQLVWARGMEA-GAVQVTPMIDGVWYLANPGDVIQALDAQGTDLIWEHRRQLPAVA 122
Db 83 VKDILGLAMSYNLESTRGVEATPVVDGIMYVSASWSVVAHIDRTGMRIV----- 132

RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF176640; MAG09249.1; .
 DR HSSP; Q46444; 1XB0.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. .; IEA.
 DR GO; GO:0005489; F:electron transporter activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001479; Bac PQO.
 DR InterPro; IPR000345; CytC_heme_B8.
 DR InterPro; IPR009056; Cytochrome_C.
 DR InterPro; IPR003088; Cyt_C1.
 DR InterPro; IPR002372; PQO repeat.
 DR InterPro; IPR011047; Quin_ald_DH_like.
 DR Pfam; PF000034; Cytochrom_C_1.
 DR Pfam; PF01011; PQO; 3.
 DR SMART; SM00564; PQO; 5.
 DR PROSITE; PS00364; BACTERIAL_PQO_2; 1.
 DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN 1.
 DR SEQUENCE 695 AA; 75842 MW; D29A698A642E1A83 CRC64;
 Query Match 19.4%; Score 594; DB 2; Length 695;
 Best Local Similarity 30.7%; Pred. No. 2,9e-32;
 Matches 193; Conservative 87; Mismatches 235; Indels 114; Gaps 30;
 QY 6 LKRTSAVLLTLTA--PAFAQVTPITDEL--ANPAGWINYGRNENYRHSPLTQITA 61
 DB 4 LTQSRLVCLTLTLVAAGAAKVDAAIRASBOOSELSHGRTYAEGRFPLQIDA 63
 QY 62 DNVGQLQVMA-----RGMEAGAVQVTPMTHDGMVTLAN--GVVQALD--AQTGLIWEHR 115
 DB 64 GNVGKGLAVYLDLENNRGLAE-----TPVSDGVYASLSWSRVAVDILRSKRLLMO-- 116
 QY 116 RQLPAVATLNAQGR-----KRGVALYGTSLYFSSWDMHLIALDMETGVVPEV 164
 DB 117 -----FPQVDRGHSRTCCDAVNRGVALMKGKTYVALDGRLLADAKIGRELMS- 167
 QY 165 ERSGGEDLTSTNTGPIVANG--VIVAGSTCQYSPYCGFISGHDSATGEELMRNHPF--- 220
 DB 168 EORTDPAPKYSITGAPVIVKGYIIGNGALEYVGRG--FESAYDAETGKNAMRFYTPGDB 226
 QY 221 -ORGE-----EGDETMGNDFEARMNG-----VMGQITYPVTNLVYVSGSTGVPASETOR 270
 DB 227 AQPTEHELAEAAKTKGD--QYWKUGGGGTWDSMAVDELDLTYTG--TGNGSPNNREI 283
 QY 271 GTRGG--TLVGTNTRFAVRPDTEIYWRHQTLPNDWDOCTFEMVAVAVDVQPSALEMEG 328
 DB 284 RSPGGGDMLY--LSIILARPDGSKLMLHNYQTTEGFTMDFTAQITLTL-----ELDG 336
 QY 329 LRALNPAATGERRVLTGAPCKTGTMSPFASGEFLMARDNYTMIASID-ETGL-VT 386
 DB 337 -----KPKRVLMQAP--KNGFFVLDRAATGELLSAEKFGKVTMAEKVDLATGRPYE 385
 QY 387 VNEDAVLKELDVEYDVCPTEFLGGRDWSAALNPDGTGIFLPLNNAICYDIMAADOE-----F 442
 DB 386 VPSSRREKEGVVWVW---PSSFGAHNMHSMFNQGTLMKTIPTG---ELPGYIRNEGATF 438
 QY 443 SALDVNTSATKLAQGFENMGRIIDALDISTGRTLSAER-----PAANT--SEVLST 493
 DB 439 KKIDGNTGT-----GFSPTHEIPRDVAGSALLAMDPRKREAWEPHSPFVNGGTLSST 492
 QY 494 AGGVVNGGDRYERALSQETGELTQWQARLATVATGQAISYELDGVQYAI-----GAAG 548
 DB 493 AGNLVQGGTRDGGDLHAYSAKQRLMSFAAQGTIVAAPISFSDGEQYVAVNAGMGGAAP 552
 QY 549 LTYGQLNAPLEAIDSTSVGNAIYFAL 577
 DB 553 LIGG---DALADGVNRLS---RLVEXL 575
 RESULT 13
 Q89GY2 PRELIMINARY; PRT; 602 AA.
 AC Q89GY2;

DT 01-UTN-2003 (TREMBLrel. 24, Created)
 DT 01-UTN-2003 (TREMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Methanol dehydrogenase large subunit-like protein.
 GN Name=mxaf; Order=locusNames=blt6213;
 OS Bradyrhizobium japonicum.
 OC Bradyrhizobium japonicum.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Bradyrhizobiaceae; Bradyrhizobium.
 RX MEDLINE:22464398; PubMed:12597275;
 RA Kaneo T., Nakamura Y., Sato S., Minamisaawa K., Uehimi T.,
 RA Sasamoto S., Watanabe A., Ideasa K., Iizuchi M., Kawashima K.,
 RA Kohara M., Matsunoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,
 RA Tabata S.;
 RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
 RT Bradyrhizobium japonicum USDA110."
 RL DNA Res 9:189-197(2002).
 DR EMBL; AF005957; BAC51478.1; .
 DR HSSP; P38539; 1G72.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR002372; PQO repeat.
 DR InterPro; IPR011047; Quin_ald_DH_like.
 DR Pfam; PF01011; PQO; 3.
 DR SMART; SM00564; PQO; 6.
 KW Complete proteome.
 SQ SEQUENCE 602 AA; 64779 MW; 96026018D3F7EE9D CRC64;
 Query Match 19.3%; Score 593.5; DB 2; Length 602;
 Best Local Similarity 28.0%; Pred. No. 2,6e-32;
 Matches 181; Conservative 110; Mismatches 230; Indels 125; Gaps 26;
 QY 5 TLRTSAVLLTLTAPAFQVTPITDEL--ANPAGWINYGRNENYRHSPLTQITA 61
 DB 8 TYLGSAAALAVGSASA-----NDELIRKXQNPK--DWMPAGADVANTRYSKLQINA 57
 QY 62 DNVGQLQVMA-----RGMEAGAVQVTPMTHDGMVTLAN--GVVQALD--AQTGLIWEHR 113
 DB 58 QNVGQLQVAMTFSTGVLRGHEG---PLTIGNMVTHVTPFNKVAITDLSNNKTKVWK 112
 QY 114 HR-RQLPAVATLNAQGRKRGVALYGTSLYFSSWDMHLIALDMETGVVPEVVERSGEDG 172
 DB 113 YEPKQDPVIVPVMCCDVTNRLGSLYGDGKILHQADVTNLVALDATGVQVMSATNGDSKG 172
 QY 173 LFSNTGTVANGYIVAGSTCQYSPYCGFISGHDSATGEELMRNHPFQPEE----- 225
 DB 173 QTGTSALVAVKDYLVGISGEFEG--VOCHVTAYDLKSGKQVWR--AFSEGPDDQIKVDPAK 230
 QY 226 -----GDETWGNDFEARMNG-----VMGQITYPVTNLVYVSGSTGVPASETOR 270
 DB 221 TTSLGKPVGADSLKTKWGD---QMKUGGGCTKGMVYDPAALNVYVGS--NSTWNPK 285
 QY 271 GTRGGTTLVGTNTRFAVRPDTEIYWRHQTLPNDWDOCTFEMVAVAVDVQPSALEMEGLR 330
 DB 286 QRPQGNKK--SMITFARBADITGMAXKVYQMTPEHDMEDVGVNEMILSQOQINGA----- 338
 QY 331 AINPAAATGERRVLTGAPCKTGTMSPFASGEFLMARDNYTMIASID--DTNNTMI----- 376
 DB 339 -----RKLITHTD--RNGLGVTMDRESELIVAEKIDPKKNWISGVDMKNSPTYG 387
 QY 377 -----ASIDETGLVTNEDAVLKELDVEYDVCPTEFLGGRDWSAALNPDGTGIFLPLN 429
 DB 388 RPKYLDAASTDKAG-----EDHNK-----GICPALGTQDEQAPAYSPDTQFVYPTN 436
 QY 430 NACVDIMAVDSEFALDVNTSATKLAAP--GFENMRIDAIDSTGRTLSMRAPRANY 487
 DB 437 HVCWDYEPFKVSYTAGQPT--VGATLSYTPROGESHMGNFTAMDKTIAIVSNKEQFSVW 495
 QY 488 SPVLSTAGGVVNGGTRYERALSQETGELTQWQARLATVATGQAISYELDGVQYAI--G 545
 DB 496 SGALATAGGVVFTLGGYLKAVDAKSGKELYKTPSGIIGNVTYENGGKQVAVAVLSG 555

QY 546 AGGLTGTCTLNAPL-----AAIDSTSVGNATYFALP 578
DB 556 VGGMA-GTGLAAGLDTPTAGLAAGVYALSNYTHALGTLTYFSLP 600

RESULT 14

Q6MTW7 PRELIMINARY; PRT; 588 AA.

AC 06MTW7;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE Alcohol dehydrogenase (Acceptor).
GN CEFNames-RNGR00479;
OS Rhizobium sp. (strain NGR234).
OC plasmid megaplasmid 2.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_TaxID=394;
RN SEQUENCE FROM N.A.
RC STRAIN=NGR234;
RA Broughton W.J., Perret X., Staehelin C., Schmitz R.A., Raasch C.,
RA Liesegang H., Gottschalk G., Streite W.R., to the EMBL/GenBank/DBJ databases.
RL SDBL: AY316747; AA087251.1;
DR Interpro: IPR002372; PQQ_repeat.
DR Interpro: IPR011047; Quin_alc_DH_like.
DR Pfam: PF01011; PQQ; 2.
DR SMART; SM00564; PQQ; 7.
KM Plasmid.
SQ SEQUENCE 588 AA; 63953 MW; 5A6007CC5B3F6329 CRC64;

Query Match 19.0%; Score 582; DB 2; Length 588;
Best Local Similarity 27.2%; Pred. No. 1.6e-31;
Matches 174; Conservative 109; Mismatches 229; Indels 128; Gaps 26;

QY 7 LRTSAVLLITPAAPAOVTPITDELLAN--PRAGEWINGRNOENYRHSPLTQITADNV 64
DB 4 LMTAVTLLSLAGLASAEV--TEEDLAKDATTGVLVTGMGRGIGQRYSPLFTLAKTV 60
QY 65 GOLQVWA-----RGMEGAVQVTPMIDGVMTLANPGDVIQALDAQTGLIWEHRR 116
DB 61 KNLLPAMGFSLSGKRGQGES-----QPIYDGVMTITASYSRLVAIDIKTGKELWQFDA 115
QY 117 QLPRAVATLNAQGRKRGVALYGTSLYFSSWDNHLIALDMETGOVVDVVRGSGEDGLTSN 176
DB 116 RLPE-GILPCCDVINRGAATYGDNIYFGTLDARLVALNRKTGDVWNNKIANTKEGY-SY 173
QY 177 TTGPIVANGVIVAGSTCOYSPYGC--ISGHSATGEELW-----RNHFLPQGEEG-- 226
DB 174 TAAPLVNGLITLINS--GGEFGIVGEVQARDAKTGELVWTRPVEIGHVGSYKXOSSMT 231
QY 227 ---DETGNDFEARMTG---VWGQITYPVTNIVFGSTGVP-----ASETORGTP 273
DB 232 GTLNATVPGDL---WKTGGATWLGSSYDADTDLVLTGTCNPAWNSHLRNAGKPVENGK 288
QY 274 GGTLYGTNTPAVRPDGEIVWRHQTLPRDNWDECTFEMMVANVDVOPSAEHEGLRAIN 333
DB 289 GDNLYAA-SRLGINPENGELIKMHQTLPRGMDPDGVNE-----VVFPSKDGTRF- 339
QY 334 PNAATGERRVLTGAPCKTGTWMSFDAAGEFL-----WARDNTYNTMIASIDETG 383
DB 340 ---ATAADR-----NGPFYVLRNEDGKFAVAPVKNIWAK-----GIDETG 378
QY 384 LVTVNEADVILKELDEVY-----VCPFLIGRDMSSAALNPDGIYELPLNNACYDI 435
DB 379 RPIYAENRPDDBSSAADGKGQGVSVPSFLGKMMMPAAYSQKSELFVPSNE----- 433
QY 436 MAVDOESALDVNTSATAX-----LAPGFEN-NGRIDAIDISTGRTILMSARP 483
DB 434 ---WGDINWEPISYKKGAAYLGAGFTIKPLFEDYISGLKALDPNKGKIMKWEYKNA 486

QY 484 AANYSPVLSAGGVVNGTDRYFRALSDFTGTLMOARLATVATQALSYELDVOYIA 543
DB 487 APLMGVATAGALIVFGTPEGEFFKALDDDETGEELMSFGVSGGVQPTTWEDGGQYVS 546
QY 544 I--GAGGLT--YGMQUNAPLAEALDSVGNATYFALPQ 579
DB 547 VISGGMGAVPLMKGB-----YAKKVTYLNQGMWWTFRLRK 582

RESULT 15

AA087251 PRELIMINARY; PRT; 588 AA.

AC AA087251;
DT 02-MAR-2004 (T-EMBLrel. 27, Created)
DT 02-MAR-2004 (T-EMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (T-EMBLrel. 27, Last annotation update)
DE Alcohol dehydrogenase (Acceptor).
GN RNGR00479;
OS Rhizobium sp. (strain NGR234).
OC plasmid megaplasmid 2.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_TaxID=394;
RN SEQUENCE FROM N.A.
RC STRAIN=NGR234;
RA Broughton W.J., Perret X., Staehelin C., Schmitz R.A., Raasch C.,
RA Liesegang H., Gottschalk G., Streite W.R., to the EMBL/GenBank/DBJ databases.
RL SDBL: AY316747; AA087251.1;
DR Interpro: IPR002372; PQQ_repeat.
DR Interpro: IPR011047; Quin_alc_DH_like.
DR Pfam: PF01011; PQQ; 2.
DR SMART; SM00564; PQQ; 7.
KM Plasmid.
SQ SEQUENCE 588 AA; 63953 MW; 5A6007CC5B3F6329 CRC64;

Query Match 19.0%; Score 582; DB 2; Length 588;
Best Local Similarity 27.2%; Pred. No. 1.6e-31;
Matches 174; Conservative 109; Mismatches 229; Indels 128; Gaps 26;

QY 7 LRTSAVLLITPAAPAOVTPITDELLAN--PRAGEWINGRNOENYRHSPLTQITADNV 64
DB 4 LMTAVTLLSLAGLASAEV--TEEDLAKDATTGVLVTGMGRGIGQRYSPLFTLAKTV 60
QY 65 GOLQVWA-----RGMEGAVQVTPMIDGVMTLANPGDVIQALDAQTGLIWEHRR 116
DB 61 KNLLPAMGFSLSGKRGQGES-----QPIYDGVMTITASYSRLVAIDIKTGKELWQFDA 115
QY 117 QLPRAVATLNAQGRKRGVALYGTSLYFSSWDNHLIALDMETGOVVDVVRGSGEDGLTSN 176
DB 116 RLPE-GILPCCDVINRGAATYGDNIYFGTLDARLVALNRKTGDVWNNKIANTKEGY-SY 173
QY 177 TTGPIVANGVIVAGSTCOYSPYGC--ISGHSATGEELW-----RNHFLPQGEEG-- 226
DB 174 TAAPLVNGLITLINS--GGEFGIVGEVQARDAKTGELVWTRPVEIGHVGSYKXOSSMT 231
QY 227 ---DETGNDFEARMTG---VWGQITYPVTNIVFGSTGVP-----ASETORGTP 273
DB 232 GTLNATVPGDL---WKTGGATWLGSSYDADTDLVLTGTCNPAWNSHLRNAGKPVENGK 288
QY 274 GGTLYGTNTPAVRPDGEIVWRHQTLPRDNWDECTFEMMVANVDVOPSAEHEGLRAIN 333
DB 289 GDNLYAA-SRLGINPENGELIKMHQTLPRGMDPDGVNE-----VVFPSKDGTRF- 339
QY 334 PNAATGERRVLTGAPCKTGTWMSFDAAGEFL-----WARDNTYNTMIASIDETG 383
DB 340 ---ATAADR-----NGPFYVLRNEDGKFAVAPVKNIWAK-----GIDETG 378
QY 384 LVTVNEADVILKELDEVY-----VCPFLIGRDMSSAALNPDGIYELPLNNACYDI 435
DB 379 RPIYAENRPDDBSSAADGKGQGVSVPSFLGKMMMPAAYSQKSELFVPSNE----- 433
QY 436 MAVDOESALDVNTSATAX-----LAPGFEN-NGRIDAIDISTGRTILMSARP 483

Thu Nov 18 06:35:34 2004

us-10-802-682-8.rup

Page 11

Db 434 -----WGDINDINNEPISYKKAHAHAGCFTTKLPEDFYISLKAIDPNNGEITMEYKCA 486

OY 484 AANISPVLSLAGVNVNGSTDRYFRALSOETGETIMQARLATAVGAISYELDGVVYA 543

Db 487 APLMGVMAVAGGLVVGTPPEGEFRALDELTEBEELMSQTSQSGVAGQPTWEDQGEQYVS 546

OY 544 I--GAGGLT--YGTOLNAPLAEIISTSGNAINYFALPO 579

Db 547 VISMGGAAPVLMNGE---VAKKVATILNMGGMVVFTRLPK 582

Search completed: November 18, 2004, 05:19:39
Job time : 209 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: November 18, 2004, 04:39:43 ; Search time 42 Seconds

(without alignments)
1326.417 Million cell updates/sec

Title: US-10-802-682-8

Perfect score: 3069

Sequence: 1 MNPTTLRTSAVLLTAPA.....AEADSTSVGNALYFPALPO 579

Scoring table: BLOSUM62

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Database: 1: PIR.79:*

2: PIR.1:*

3: PIR.3:*

4: PIR.4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	597	19.5	708	2	SS2317
2	568	18.5	601	2	E95863
3	552	18.0	623	2	B83599
4	551.5	18.0	738	2	S14270
5	535.5	17.4	626	2	JQ0706
6	511.5	16.7	742	2	A49340
7	491.5	16.0	742	2	JS0326
8	435.5	14.2	573	2	S68591
9	399	13.0	796	1	JV0107
10	399	13.0	796	2	H85495
11	399	13.0	796	2	H90644
12	398.5	13.0	796	2	AG0523
13	375.5	12.2	801	1	S00943
14	374.5	12.2	809	2	B98314
15	374.5	12.2	809	2	AI2968
16	351.5	11.5	808	1	QPKEX
17	342.5	11.2	778	2	G98221
18	342.5	11.2	778	2	AI3064
19	341.5	11.1	803	2	F83360
20	335	10.9	809	2	A55547
21	278	9.1	639	2	JC4881
22	224	7.3	221	2	A41378
23	172.5	5.6	407	2	H69064
24	168	5.5	524	2	A82580
25	159.5	5.2	407	2	C82804
26	152.5	5.0	668	2	C75264
27	151.5	4.9	839	2	D97013
28	144.5	4.7	613	2	F69424
29	140.5	4.6	3029	2	S76109

30	139	4.5	827	2	F64512	hypothetical prote
31	136	4.4	392	2	F91050	probable dehydroge
32	136	4.4	392	2	B65895	probable dehydroge
33	135	4.4	603	2	F72237	conserved hypotnet
34	133	4.3	392	2	G65027	hypothetical prote
35	132	4.3	2535	2	AC0304	probable hemolysin
36	131.5	4.3	424	2	T29127	hypothetical prote
37	130.5	4.3	1197	2	D82696	hypothetical prote
38	129.5	4.2	386	2	A82284	conserved hypotnet
39	128.5	4.2	2554	2	A83528	extracellular seri
40	128	4.2	1939	2	A82018	hypothetical prote
41	127.5	4.2	393	2	AD0350	probable lipoprote
42	127	4.1	392	2	AG0821	probable lipoprote
43	126	4.1	856	2	T00349	Avicelase III - As
44	126	4.1	946	2	F84280	Arp-dependent heli
45	124	4.0	796	2	T39962	coatomer complex b

ALIGNMENTS

RESULT 1	SS2317	quinohemoprotein ethanol dehydrogenase (EC 1.1.99.-) type 1 precursor - Comamonas testost
C/Species:	Comamonas testostestoni	
C/Date:	08-May-1995	#sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C/Accession:	S62366; S62373; S65908; S52317	
R/Stoortvogel, J.; Kraayveld, D.E.; van Slufts, C.A.; Jongejan, J.A.; de Vries, S.; Duine, Eur. J. Biochem. 235, 690-698, 1996		
A/Title:	Characterization of the gene encoding quinohemoprotein ethanol dehydrogenase of	
A/Reference number:	S62366; MUID:96184549; PMID:865419	
A/Accession:	S62366	
A/Status:	preliminary	
A/Molecule type:	DNA	
A/Residues:	1-708 <ST01>	
A/Cross-references:	UNIPROT:Q46444; EMBL:X81880; NID:g663195; PIDN:CA57464.1; PID:g66311	
A/Accession:	S62373	
A/Status:	preliminary	
A/Molecule type:	protein	
A/Residues:	32-54 <ST02>	
Ride Jong, G.A.H.; Geertlof, A.; Stoortvogel, J.; Jongejan, J.A.; de Vries, S.; Duine, Eur. J. Biochem. 230, 899-905, 1995		
A/Title:	Quinohemoprotein ethanol dehydrogenase from Comamonas testostestoni. Purification	
A/Reference number:	S65908; MUID:95324580; PMID:7601151	
A/Accession:	S65908	
A/Molecule type:	protein	
A/Residues:	32-50, 'X', 52-54; 477-483, 'X', 485-490 <DEJ>	
A/Experimental source:	ATCC 15667	
C/Genetics:		
A:Gene:	qhdh	
C/Superfamily:	glucose dehydrogenase (pyrroloquinoline-quinone)	
C/Keywords:	oxidoreductase, quinoprotein	
F1-31/Domain:	signal sequence #status predicted <SIG>	
F132-708/Product:	quinohemoprotein ethanol dehydrogenase type 1 #status experimental <MAV	
Query Match	19.5%; Score 597; DB 2; Length 708;	
Best Local Similarity	28.9%; Pred. No. 1.2e-33;	
Matches	178; Conservative 91; Mismatches 251; Indels 96; Gaps 21;	
QY	16	LTAPAFQVTPIT-----DELLANP-AGEKINYGNOENYRSPPLQITADN 63
DB	23	LGSAALFAQTGPAQAAAQVAVQVVDGDFIRANVARTPDMPTIGVVAETFRYSRLQIINAN 82
QY	64	VGQQLVVARGMEA-GAVQVTPMTIDGQVYLANPGDVQALDAQCTGDLIWEHRQLPVA 122
DB	83	VKDLGLANSYLSREYGVATPVVDGIMVYSAKMSVHAIDTITGNRIW-----132
QY	123	TLNAQGR-----KGVALLYGTSLYFSSWDNHLALDMETGVQVDFVERGSGED 171
DB	133	TYDPQIDSTGFGKCCDVVNRGVALMKGVYVGAWDGRLIALDAATGKRVHONTFFEQK 192
QY	172	GLTSWTGPVANGVIVAGST-CQYSPYGCFSIGHDSATGSELRNHFIP---QPGEE- 225

```

Db      193 GSLITIGAPRVFKVIGKGALEYGRC-YITAYDAETGERKMRFSVPDPSPKFEDE 251
Qy      226 ---GDETW---GNDEPARMTGVMGOITYPVTLNLFYGSIGVGPASSTOBTGCTILY 278
Db      252 SMKRAARTWDPGKMWEDAGGGTMDSMFEDALINTMYGTGNGSPMSHKVRS PKGDDL 311
Qy      279 GINTPRANVPDTGEIYWRHQTLPDRNDWDOECTFEMVNAVVDQPSAEMGLAALNPNAI 338
Db      312 YLASIALVADPTGKYMHWIETPGDNWDYTSOTPMILADIKI-----AG 355
Qy      339 GERVRVLTGAPCKTGTWMSFPAASGEFLMARDTNYTMNLASIDETGL-VTVNEDAVLKELD 397
Db      356 KPRKVLHNP-KNGFFFLVLDRTNGKFLSAKNFVPMWASGYDGHGPKIGI---AARDGS 411
Qy      398 VERDVCTPLFGDNSSAALNPDTGIYFLPNNACCDIADVDQEFALVYNTSA----- 452
Db      412 KPDAVGPYGAHNMFMPSFNPTGLVYLPQAVNPANLND-DKRM---FNQAGPKQ 466
Qy      453 ---TAKL---AGFEMWGSIDAIDISTGRTLSABRPANYSPLVSTAGVFNNG 502
Db      467 SGTGNMTAKFPNABPKSKPFGLLAMDPAQAKAAVSEHVSWMNGTLTTGNVVFQGT 526
Qy      503 TDRYFALSGETITMQLATVATGQAISYELDGVYIATGAG-GLTYGTQLNAPLAE 561
Db      527 ADGRLVAYHAATGEKLMKWEAPGTGVVAAPSTVYVADGRQYVAVAGWGYYGL---AAR 581
Qy      562 AIDSTSVGNAYVFAL 577
Db      582 ATERQPG-TVYTFVV 596

```

RESULT 2

```

B95863
alcohol dehydrogenase (acceptor) (EC 1.1.99.8) large subunit [imported] - Sinorhizobium
C:Species: Sinorhizobium meliloti
C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C:Accession: B95863
R:Finan, T.M.; Weidner, S.; Wong, K.; Bulhmerster, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb pSYM3 megaplasmid from the N2-fixing endo
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: B95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KIR>
A:Cross-references: UNIPROT:Q92WY9; GB:AL591985; EIDN:CA048573.1; PID:G15140045; GSPDB:Q
A:Experimental source: strain 1021, megaplasmid pSYM3
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
Pel, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federpiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.T.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: SMD20173
A:Genome: plasmid
C:Keywords: oxidoreductase

```

```

Query Match      18.5%; Score 568; DB 2; Length 601;
Best Local Similarity 27.7%; Pred. No. 1e-31;
Matches 173; Conservative 101; Mismatches 243; Indels 108; Gaps 24;

Qy      23 AOVTPITDEL-LANPAGEMVINGRNOENYRHSPLTQTADNVGQLVMA-----RG 74
Db      17 AOVAFANDELQKLIDDE-NQMAIQOTGDYANLRYSKLDQINKDVGKQAVAMFSTGVLRG 75
Qy      75 MEGAVQVTEMHIDGMYNLNP-CDVYQALD-AQTGDLIMEHR-RLPVAATLNAAGDGR 131
Db      76 HEG-----SPLVIGDLMTYHTFPFNPNTVYALDSKQDQITWKIETPKQDDPNVIVMCDTVN 130

```

```

Qy      132 RGVLYGTSLYFSSNDNHLALDMETGOVVDYERGGEGDLTNTTGP-VANGVIAGS 191
Db      131 RGVAYADNKKLPADQDTTVALDLKTKVIMSVKNGDATGENTATVMEVKKIIVGIS 190
Qy      192 TCQSPYGCFLSGHDSATGELWRNHPI-----PQGEDEFTWAND 233
Db      191 GGEFGVNG-HYTAISMADKVLWRGVSMGPDSDLIDPEKTHLGRKVGKDSGLTTWEGD 249
Qy      234 FEARMTG---VQCITYPVTLNLFYGSIGVGPASSTOBTGCTILYNTTFAVRPDT 290
Db      250 ---CWKIGGGTGWGWSYDPEENLVYGTGNPSTWNTQS--PDNRM-SYTFARDVDT 303
Qy      291 GEIYWRHQTLPDRNDWDOECTFEMVNAVVDQPSAEMGLAALNPNAATGERRVLTGA PCK 350
Db      304 GMAKVLQMPHDEWDVDGVNEMILTBQIDGK-----DRKILTFED-R 346
Qy      351 TGTWMSDPAASGEFL-----WARDNTYTMNLASIDETGLVYNEDAVLKEIDVEY 400
Db      347 NGFQYMDRVYTGELVAEKYDPTVNMATE---VVMPEKSDKYG---RPQVAQYSTEQ 398
Qy      401 D-----VCPFLGDRWSSAALNPDTGIYFLPNNACCDIADVDQEFALVYNTSA 452
Db      399 NGEDTNTTGVCPALGTRKQDQPAAYSPKTELFPYPTNHVCMDEPFRVSYTAGQPY-VGA 457
Qy      453 TAKLAPGEN---MGRIDAIDISTGRTLSABRPANYSPLVSTAGVFNNGTDRYFRA 509
Db      458 TLSMYPPKDSHGNGNFIANDNKEGKIKWSLPBFSVMSALATADGVFFYGLSGFLKA 517
Qy      510 LSGETITMQLATVATGQAISYELDGVYIATGAG-GLTYGTQLNAPLAEIDS-- 565
Db      518 VDAATGKELRFKTPSGVIGNMTYAREGKQYVAVASGVGMA-GIGLAAGLNPITREGIG 576
Qy      566 -----TSVGNAYVFALPQ 579
Db      577 AVGVASLSNYTALGTLTFVKLPE 601

```

RESULT 3

```

B83399
guiniprotein alcohol dehydrogenase PA1982 [imported] - Pseudomonas aeruginosa (strain PAC
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: B83399
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mironuchi, S.D.; Warrenner, P.; Hickey, M.U.; Bri
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: B83399
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-623 <STO>
A:Cross-references: UNIPROT:Q924U7; GB:A5004624; GB:A5004091; NID:G9947973; PIDN:AAG0537C
A:Experimental source: strain PA01
C:Genetics:
A:Gene: exah; PA1982

```

```

Query Match      18.0%; Score 552; DB 2; Length 623;
Best Local Similarity 27.2%; Pred. No. 1.4e-30;
Matches 180; Conservative 110; Mismatches 220; Indels 142; Gaps 28;

Qy      3 PTTLLRTSA-----AVLLTAPAFAOVTPITDELLAN--PPAGEMVINGRNOENYRHS 55
Db      9 PAGLLRPSLHCLAFVALGSAALAK--DVTWEDINDKXTGDDVLYQMGTHAQRWSP 66
Qy      56 LQTITADNVGQLVMA-----RQMEKAGAVQVTEMHIDGMYNLNPBGDVYQALDAQT 107
Db      67 LQVYADVVFELTFAWSYSPGDEKQROESQAI-----VBDGYIYVYASVSRFLPDAK 121
Qy      108 GDILWEHRRL-----PAVATLNAQGRKGVVALYGTSLYFSSNDNHLALDMETGOVVD 162
Db      122 GKRLMTYHRLPDDIRFCQVYV-----RGAATYGRKVFPGTLDASVVALNKVTKGVVW 175

```



```

QY 163 DVERGSGEDGLTNTTGPPIYANG-----VIVAGST-CQVSPYCCFISGHSATGEELMEN 216
DB 176 KKKFADHAGAYTWTGAPTIIYKDKTKGKVLIIHSSSGDEFGVGRIPA-RPDTGEEIWMR 234
QY 217 HFIPOPEEG-----DETMDNDEFA-----RMWTG---VWGQITYPVTN 253
DB 235 PFI--EEMKRLKSGSTVTGDKAPBSPPDRNSPTGKESMSHGGAPOASFDATN 292
QY 254 LVRYGSTGVPASETORGTGEG-----TLYGNTRFAPAPDTGEIWMHQTLPRDNMQ 307
DB 293 TIIVGAGNPGRMTWARTAKGNPHDYSLY-TSGQVNDPSSGEVWKFYQHTPNDAWF 351
QY 308 ECFEEMVAVAVDQPSAEMEGLAIPNATGERRVLTGAPCKTGTMTSADAAGE----- 363
DB 352 SGNNELV-----FDYKADGKIYKATAHADNKGFFVVDNRSGHLQNA 395
QY 364 -----FLMARDTNNTMTIASID-ETGLVTNED-AVLKELDVEY-----DVCPTLGR 410
DB 396 FPFVDNITWA-----SHIDLKTRGPVERREGORPLPEPGQKHGKAVESSPPLGKR 446
QY 411 DMSAALNPRTGIYELPANNACDIMAVDQESALDVYNTSATKLAPEF-----ENM 463
DB 447 NMPPMVYSQGTGFYYPANH-----WKEDYITEEVSYTKGSAYIGMGFRIKMYDHY 499
QY 464 GRIDAIDISTGRTIWSAERPANYSPVLSTAGGVFNCGTDRYFRALSOETGELMQAL 523
DB 500 GSRANDPVSGKVMHEKHLPLMAGVLATAGNLVFTGTGDYFKAPDAKSGELMKPOT 559
QY 524 ATATGAGTISYEIDGVOY-IAIGAGLT--YG--TQNLAPAEIADISTVGNALYVA 576
DB 560 GSGIVSPPTWEDGGEQYLGVTVGAGVPLMGDMADLTRPVAQ-----GGSFWVK 612
QY 577 LP 578
DB 613 LP 614

```

RESULT 4
S14270
alcohol dehydrogenase, membrane-bound (BC 1.1.1.-) 72K chain precursor - Acetobacter pol
C:Species: Acetobacter polyoxogenes
C:Accession: S14270
C:Date: 21-Nov-1993 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
R:Tamaki, T.; Fukaya, M.; Takemura, H.; Tayama, K.; Okumura, H.; Kawamura, Y.; Nishiyama
Biochim. Biophys. Acta 1088, 292-300, 1991
A:Title: Cloning and sequencing of the gene cluster encoding two subunits of membrane-bc
A:Reference number: S14270; MUID:91159482; PMID:2001402
A:Accession: S14270
A:Molecule type: DNA
A:Residues: 1-738 <TAM>
A:Cross-references: UNIPROT:P28036; GB:D00635; NID:g216185; PIDN:BAA00528.1; PID:g216186
A:Experimental source: strain NB1028
C:Complex: heterodimer of 72K and 44K chains
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
C:Keywords: alcohol metabolism; membrane protein; oxidoreductase
F:1-35/Domain: signal sequence #status predicted <SIG>
F:36-738/Product: alcohol dehydrogenase 72K chain #status predicted <MAT>

Query Match 18.0%; Score 551.5; DB 2; Length 738;
Best Local Similarity 27.0%; Pred. No. 1.9e-30;
Matches 164; Conservative 91; Mismatches 253; Indels 99; Gaps 17;

```

QY 5 TLRTSAVLLTAPAAFAQVTP--ITDELL--ANPAGWINTGRQENYRHSPLTOI 59
DB 15 TLTAGTICALLISGATVTSADGCGATGEALIHADHGMNMTYGRITSQDYSPLDOI 74
QY 60 TADNVGQLQVM-----ARGWEAGAVQYTPMHCNHWYLIAPGVYQALDQOTGLWE 113
DB 75 NNSNVGNLKLAMYLDDTNRGGEG---TFVLIDVMTATTNWSMKKAVDAGTKLMS 129
QY 114 HRRQLPA-VATLNAQGRKRGAVALYGTSLYFSSWMDNHLIALDMETGVVFDV----- 164

```

```

DB 130 YDPRVGNIAKGCDDTYNRGAAYWNGKYPGTDFDRLIALDAKGLWMSVNTI-PPEAE 189
QY 165 ---ERGSGEDGLTNTTGPPIYANG-VIVAGSTCOVSPYCCFISGHSATGELMRNHP 220
DB 190 LKQKSYIVDG-----APRIAKGRVITIGNGSGSERGANG-FVSAPDADTGKVDREFTVP 242
QY 221 QGEEGD-----ETWGNDFEARMTG-----VWGQITYPVTNLYFYG 258
DB 243 NFKNEPDAASSVLMNKAYQWTS-----PTGAWTRQCGCGTWDSTIVYDPAVLVLG 295
QY 259 STGVGPASETORGTGCGTLYGNTRFAVRPTGEIWMHQTLPRDNMQDETFEEMVAVY 316
DB 296 VNGSPFNMYKTRSGEKNLNLGSLVHLKPEGTGYWHFQETPMDOQMFSDQITMLDL 355
QY 319 DVQPSAEMEGLAIPNATGERRVLTGAPCKTGTMTSFDAASEFLMARDTNNTMTIAS 378
DB 356 FI-----NGETRHVIVHARKNGPFYIDAKTGEBTSGKYVYVNMASG 398
QY 379 ID-ETGLVTNEDAVLKELDVEYDVCPTFLGGRMSSAALNPDTGIYELPANNACY--- 433
DB 399 LDPKTRGPIYNDALYTLTGKEWYGPEDLGHNFPAAAFSPKGTGLYIIPAQVFFLYTN 458
QY 434 -----DIMAYDOESALDVYNTSATKLAPEFENMGRIADAIDISTGRTIWSAERPA 484
DB 459 QVCGFTPHPDGSMNIGLDMNKVGIID-SPEAKQAFVKDLKGIIVAMDQKQAEARVPHKG 517
QY 485 ANYSVVLSTAGGVFNCGTDRYFRALSOETGELTQMARLATVATGQALSYELDGVOYTA 543
DB 518 FPNQGIATGDDLLFOGLANGEFHAYDATNGSDLFHFAADSGIITAPVYTLANGKQYAV 577
QY 544 -IGAGLT 549
DB 578 EVWGCGI 584

```

RESULT 5
J00706
alcohol dehydrogenase (acceptor) (BC 1.1.1.99.8) alpha chain precursor - Methylobacterium
M:Alternate names: methanol dehydrogenase 62K large chain
C:Species: Methylobacterium extorquens
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C:Accession: J00706; S07908
R:Anderson, D.J.; Morris, C.J.; Nunn, D.N.; Anthony, C.; Lidstrom, M.E.
Gene 90, 173-176, 1990
A:Title: Nucleotide sequence of the Methylobacterium extorquens AM1 moxP and moxJ genes
A:Reference number: J00706; MUID:90373742; PMID:2116368
A:Accession: J00706
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-626 <AND>
A:Cross-references: UNIPROT:P16027; GB:M31108; NID:g150017; PIDN:AAA25380.1; PID:g150018
A:Experimental source: strain AM1
R:Nunn, D.N.; Day, D.; Anthony, C.
Biochem. J. 260, 857-862, 1989
A:Title: The second subunit of methanol dehydrogenase of Methylobacterium extorquens AM1.
A:Reference number: S04644; MUID:89350892; PMID:2504152
A:Accession: S07908
A:Molecule type: protein
A:Residues: 28-50, 'XX', 53 <NUN>
A:Note: the source is designated as Methylobacterium extorquens AM1
C:Comment: This enzyme oxidizes methanol to formaldehyde.
C:Genetics:
A:Gene: moxP
C:Keywords: alcohol metabolism; oxidoreductase
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-626/Product: alcohol dehydrogenase (acceptor) #status experimental <MAT>

Query Match 17.4%; Score 535.5; DB 2; Length 626;
Best Local Similarity 26.7%; Pred. No. 2e-29;
Matches 168; Conservative 106; Mismatches 264; Indels 91; Gaps 22;

```

QY 10 SAAVLLTAPAAFAQVTPITDELLANPAGE-WINYGRQENYRHSPLTOITADNVGQLQ 68

```


A:Reference number: I41226; MUID:93123180; PMID:8419307
 A:Accession: I41228
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-148; 'H', 150, 'KRCCHT', 157-192, 'K', 194-665, 'H', 667-796 <RES>
 A:Cross-references: GB:D12651, NID:g216555; PIDN:BA02174.1; PID:g216556
 C:Genetics:
 A:Gene: gcd
 A:Map position: 3 min
 C:Function:
 A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone
 C:Superfamily: respiratory chain
 A:Pathway: respiratory chain
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
 C:Keywords: oxidoreductase; pyrroloquinoline-quinone; respiratory chain; transmembrane F
 F/11-37/Domain: transmembrane #status predicted <TM1>
 F/41-59/Domain: transmembrane #status predicted <TM2>
 F/63-81/Domain: transmembrane #status predicted <TM3>
 F/86-110/Domain: transmembrane #status predicted <TM4>
 F/120-140/Domain: transmembrane #status predicted <TM5>
 F/93-95/Binding site: ubiquinone (Arg, Asp) #status predicted
 F/466/Active site: Asp #status predicted

Query Match 13.0%; Score 399; DB 1; Length 796;
 Best Local Similarity 24.7%; Pred. No. 8.7e-20;
 Matches 172; Conservative 91; Mismatches 247; Indels 186; Gaps 31;
 Db 150 LSAADATPAEALISPVADQ-----DMPAYGRNQEGQRFSPKQINADVNHNLKEAWVR 201
 Qy 16 LTAFAFAQ-VTPITDELANPPAGEWINGRQENYRHSPLTQTADNVGQLQVYA-- 72
 Db 150 LSAADATPAEALISPVADQ-----DMPAYGRNQEGQRFSPKQINADVNHNLKEAWVR 201
 Qy 73 -----RGMKAGAV--QVTPEIHGVMYLANPGGVIALDAQGTDLWEHRRQLPAVATLN 125
 Db 202 TGDVQKPNDEPGEITNEVTPIKVGDTLYLCTAHQRLPALDAASGKEMHNDPELKTNESEF 260
 Qy 126 AQGDRKRGVALYGTSLYFS-----NDNHLIALDMETGOV-----VF 162
 Db 261 -QHTCGVSYHKAKEATASPEVWADCPRIILPVNDGRIALNAENGKLCETPANKGVL 319
 Qy 163 DVERG--SGEDGLTSNTTGIYVANGVIV-AGS--TCQYS--PYCCFISGHSATGEBLWEN 216
 Db 320 NLQSNMPTDKRGLEYEPSPILITDKTIWAGSVTDNFTSTRETSVGVRGPDVNTGELLMA- 378
 Qy 217 HFIPO-----PGEEDETWGNDFEARMWTGWGQITTPYTNLVFYSGTGVPASETG 269
 Db 379 -FDPGADPNALPDEDETFENS-----PNSWAPAAVYAKUDLVYLPWGVTTPDIWGG 430
 Qy 270 RGTPEGTLTGNTREFAVRPDTGEIIVWRHQTLPNDWQDCTEFMMVYANDVQSAEMEG 329
 Db 431 NRTPEQERYASSI-LALNATGKLAWSYQTVHHDLMDDLPAPPTLADITV----- 480
 Qy 330 RAINPNAATGERAVLTGAPCKTGTMSPDAASGEFL-----WADDTNY----- 372
 Db 481 -----NGQKVVIYAPAKTGNIFVLDNRNGELVVPAPKEKVPQGAAGKGVVTFPTQPF 532
 Qy 373 -----TNMIASIDETGLVTNEDAVLKEI--DVEYD-----VCPTEFLGGRD 411
 Db 533 SELSFRPTKDLGSDAMWG-ATMPDQIVCRVWFHQMYEGIFTPPSEGGTLVFGNIGMFE 591
 Qy 412 WSSAALNPDTGIYF-----LPLNNACY-----DIWVDOFSALDVYNTSATA 454
 Db 592 WGGISVDPNREVALIANPMALPFVSKLIPRGPNMEQPKAKGTGIESGLOPQGVYGV 651
 Qy 455 KLAPGFENM-----GRIDAIDISTGRTLSA-----ERPANYSPV----- 490
 Db 652 TLNPFLLSPGLPCKOPAMGYISALDLKTNEVWKKRIGTPQDSMPFMPVPFPMNGMPM 711
 Qy 491 ----LSTAGGVFNNGT--DRYFRALSOETGETLMQALATVATGQA--ISYELDGVGYIA 543
 Db 712 LGGPISITAGNVLFIAATADNVYLRAYNMSNGEKLWQGLP--AGGQATPMTYEVNGKQYV 769
 Qy 544 IGAGGL--TYGTQLNAPLAEAIDSTVGNATYVFALP 578
 Db 770 ISAGHGSGFTKX-----GDYIVAYALP 792

RESULT 10
 H85495
 glucose dehydrogenase [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #ext change 09-Jul-2004
 C:Accession: H85495
 R:Perera, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
 Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimmlanta, E.; Potamouists, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: H85495
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-796 <STO>
 A:Cross-references: UNIROT:Q8946; GB:AE005174; NID:G12512839; PIDN:AA654428.1; GSPDB:G
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: gcd
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 13.0%; Score 399; DB 2; Length 796;
 Best Local Similarity 24.7%; Pred. No. 8.7e-20;
 Matches 172; Conservative 91; Mismatches 247; Indels 186; Gaps 31;
 Db 150 LSAADATPAEALISPVADQ-----DMPAYGRNQEGQRFSPKQINADVNHNLKEAWVR 201
 Qy 16 LTAFAFAQ-VTPITDELANPPAGEWINGRQENYRHSPLTQTADNVGQLQVYA-- 72
 Db 150 LSAADATPAEALISPVADQ-----DMPAYGRNQEGQRFSPKQINADVNHNLKEAWVR 201
 Qy 73 -----RGMKAGAV--QVTPEIHGVMYLANPGGVIALDAQGTDLWEHRRQLPAVATLN 125
 Db 202 TGDVQKPNDEPGEITNEVTPIKVGDTLYLCTAHQRLPALDAASGKEMHNDPELKTNESEF 260
 Qy 126 AQGDRKRGVALYGTSLYFS-----NDNHLIALDMETGOV-----VF 162
 Db 261 -QHTCGVSYHKAKEATASPEVWADCPRIILPVNDGRIALNAENGKLCETPANKGVL 319
 Qy 163 DVERG--SGEDGLTSNTTGIYVANGVIV-AGS--TCQYS--PYCCFISGHSATGEBLWEN 216
 Db 320 NLQSNMPTDKRGLEYEPSPILITDKTIWAGSVTDNFTSTRETSVGVRGPDVNTGELLMA- 378
 Qy 217 HFIPO-----PGEEDETWGNDFEARMWTGWGQITTPYTNLVFYSGTGVPASETG 269
 Db 379 -FDPGADPNALPDEDETFENS-----PNSWAPAAVYAKUDLVYLPWGVTTPDIWGG 430
 Qy 270 RGTPEGTLTGNTREFAVRPDTGEIIVWRHQTLPNDWQDCTEFMMVYANDVQSAEMEG 329
 Db 431 NRTPEQERYASSI-LALNATGKLAWSYQTVHHDLMDDLPAPPTLADITV----- 480
 Qy 330 RAINPNAATGERAVLTGAPCKTGTMSPDAASGEFL-----WADDTNY----- 372
 Db 481 -----NGQKVVIYAPAKTGNIFVLDNRNGELVVPAPKEKVPQGAAGKGVVTFPTQPF 532
 Qy 373 -----TNMIASIDETGLVTNEDAVLKEI--DVEYD-----VCPTEFLGGRD 411
 Db 533 SELSFRPTKDLGSDAMWG-ATMPDQIVCRVWFHQMYEGIFTPPSEGGTLVFGNIGMFE 591
 Qy 412 WSSAALNPDTGIYF-----LPLNNACY-----DIWVDOFSALDVYNTSATA 454
 Db 592 WGGISVDPNREVALIANPMALPFVSKLIPRGPNMEQPKAKGTGIESGLOPQGVYGV 651
 Qy 455 KLAPGFENM-----GRIDAIDISTGRTLSA-----ERPANYSPV----- 490
 Db 652 TLNPFLLSPGLPCKOPAMGYISALDLKTNEVWKKRIGTPQDSMPFMPVPFPMNGMPM 711
 Qy 491 ----LSTAGGVFNNGT--DRYFRALSOETGETLMQALATVATGQA--ISYELDGVGYIA 543
 Db 712 LGGPISITAGNVLFIAATADNVYLRAYNMSNGEKLWQGLP--AGGQATPMTYEVNGKQYV 769
 Qy 544 IGAGGL--TYGTQLNAPLAEAIDSTVGNATYVFALP 578
 Db 770 ISAGHGSGFTKX-----GDYIVAYALP 792

Db 770 ISAGHGSFGTKM-----GDYVAYALP 792

RESULT 11

H90644
glucose dehydrogenase (imported) - Escherichia coli (strain O157:H7, substrain RMD 0509
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: H90644
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhnara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A:Reference number: A9629; MUID:2115623; PMID:11258796
A:Accession: H90644
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-796 <HAV>
A:Cross-references: UNIPROT:Q8X946; GB:BA000007; PIDN:BA53351.1; PID:G13359584; GSPDB:G
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: EC50128
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 13.0%; Score 399; DB 2; Length 796;
Best Local Similarity 24.7%; Pred. No. 8,7e-20;
Matches 172; Conservative 91; Mismatches 247; Indels 186; Gaps 31;

Qy 16 LTPAAQAQ-VTPITDELANPPAGEWINGRQENYRHSPLTQTADNVGQLQVMA-- 72
Db 150 LSADATPAEAIISPVDQ-----DWPAYGRNQGGRFSPKQIADNVNHLKEAWVR 201
Qy 73 -RGEAGAV--QYTPMIDGVMTLANPGDVIOALDQGTDLWEHRRLPAYATLN 125
Db 202 TGDVYKQNDGEITNETTPIKVGDTLYLCTAHQRLPLDAASKEKMHDPKLTMSF- 260
Qy 126 AQGDRKGVALLVGTSLVTS--WDNHLIALDMETGOV-----VF 162
Db 261 -QHTVCGVSYHEAKETASPEVADCPRIILPVNDGRILAIENAGKICETFAKGVTL 319
Qy 163 DVERG--SGEDGLTSNTGPVIANGVV-AGS--TCQYS--PYGCFISGHSATGEIWMN 216
Db 320 NLQSNMDPTKPLGLEYPTSPILITDKTIWAGSVTDNSTRSTGVRGPDVNGELMA- 378
Qy 217 HFLPQ-----PGEEDETWGNDFEARMWTGVWGQITVPTNLVFGSGTVGPASETQ 269
Db 379 -FDGAKDPNAIDSDHETFTFNS-----PNSMAPAYDAKLDVYLPNGVTPPDIMWG 430
Qy 270 RGTPTGTLGTNTRFAVRPTGELVWRHQLPRDNMDQCTFEMMTANVDVQSAEMEG 329
Db 431 NRTPEQRYVASSI-LALNATGKLAWSYQTVHDLMDMLPAQPTLADIV----- 480
Qy 330 RAINPNATGERRLVLTGAPCKGTWMSFDSAAGEFL-----WARDNY- 372
Db 481 -NOKVPIVIAPAKIGNIFVLDRNGELVVPAREKVPQGAAGDVVTPQTF 532
Qy 373 -TNMIASIDETGLTVNEDAVLKL--DVEYD-----VCEFTLGSD 411
Db 533 SELSPRETKDLSGADWVG-ATMFDQVLCRVWFQMEYEGIFTPPSQGLTVFQNLGME 591
Qy 412 MSSALNPDTGIYF--LPLNACY-----DIVADQEFSLDVYNTSATA 454
Db 592 WGGISVDPENEVAILANMALPFWSKLIPRGPNMEOPKDAKTGTSBGIGQPGVPGV 651
Qy 455 KLAPEENM-----GRIDAIDISTGRTLSA-----ERPANSVP- 490
Db 652 TLNPLSPFGLPCKOPAMGYISALDLKTINEVWKRIKGTPODSMPMPMPVPVPMGMEM 711
Qy 491 -LSTAGGVVFNQGT-DRYFRALSOETGETLWQARLATVATGA--ISYELDGVYTA 543
Db 712 LGGISTAGVNLFLAATADNVLRAVYNSGEKLMQGLP--AGGQATPMTYEVNGKQYV 769
Qy 544 IGAGGL-TYGTOLNAPLAEAIIDSTVGNATVFPALP 578

Db 770 ISAGHGSFGTKM-----GDYVAYALP 792

RESULT 12

AG0523
glucose dehydrogenase (imported) - Salmonella enterica subsp. enterica serovar Typh
C:Species: Salmonella enterica subsp. enterica serovar Typh
A>Note: this species has also been called Salmonella typh
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AG0523
R:Parthill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher,
th, T.; Comerion, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moutle, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AG0523
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-796 <PAR>
A:Cross-references: GB:AU513382; PIDN:CAD01327.1; PID:G16501455; GSPDB:GN00176
C:Genetics:
A:Gene: STY0191
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 13.0%; Score 398.5; DB 2; Length 796;
Best Local Similarity 24.6%; Pred. No. 9,4e-20;
Matches 171; Conservative 96; Mismatches 235; Indels 193; Gaps 31;

Qy 19 PAFAQVPTIDELANPPAGEWINGRQENYRHSPLTQTADNVGQLQVMA----- 72
Db 156 PAATs--SIADE-----DWPAYGRNQGGRFSPKQIADNVNHLKEAWVRGDL 205
Qy 73 -RGEAGAV--QYTPMIDGVMTLANPGDVIOALDQGTDLWEHRRLPAYATLNAGD 129
Db 206 KQPNDFGRTNEVPIKVGDTLYLCTAHQRLPLDAASKEKMHDPKLTDSF--QHV 263
Qy 130 RRGVALVGTSLVTS-----WDNHLIALDMETGOV-----VF 166
Db 264 TCRGVSYHEAKADTRASPEVADCPRIILPVNDGRILPVNAETGKLCETFAKGVNLCT 323
Qy 167 GSGE--DGLTSNTGPVIAN-GVIAGS--TCQYS--PYGCFISGHSATGEIWMN 220
Db 324 NMPPTPLGLEYPTSPILITDKTIWAGSVTDNSTRSTGVRGPDVNGELMAF-- 379
Qy 221 QPGEEDETWGNDFEARMWT--GVWGQITVPTNLVFGSGTVGPASETQRTGRTGLY 278
Db 380 DPGADPNTIPADEHAFTFSPNSAPRAYDAKLDVYLPNGVTPPDIMGNKRTPEQERY 439
Qy 279 GINTFAVVRPTGELVWRHQLPRDNMDQCTFEMMTANVDVQSAEMEGRAINPNAT 338
Db 440 ASSI-LALNATGKLAWSYQTVHDLMDMLPAQPTLADIV----- 481
Qy 339 GERRVLTGAPCKGTWMSF--D-----AASGEFLMARDNYTN- 375
Db 482 GTTVVIAIAPAKTGNIFVLDRNGELVVPAREKVPQGAAGDVV-AKTOPFSLSLTPK 540
Qy 376 --IASIDETGLTVNEDAVLKL--DVEYD-----VCEFTLGSDMSSALNP 419
Db 541 KDLGADWVG-ATMFDQVLCRVWFHQLRYEGIFTPPSQGLTVFQNLGMEKGISVDP 599
Qy 420 DTGIFLPLANNAVDMAVDQEFSL-----DYNTSATXALAPGF- 460
Db 600 DRQV-----AIANPALPFWSKLIPRGPNMEOPKDAKTGTSBGIGQPGVPGV 652
Qy 461 -ENNRIDAIDISTGRTLSA-----ERPANSVP- 490
Db 653 LNPLSPFGLPCKOPAMGYISALDLKTINEVWKRIKGTPODSMPMPMPVPVPMGMEM 712
Qy 491 -LSTAGGVVFNQGT-DRYFRALSOETGETLWQARLATVATGA--ISYELDGVYTA 544

Db 713 GGPSTAGNVLFIATADNLYRAYNMSKEXLMQGRSLP--AGGQATPMTYEVNKGQYVVI 770
QY 545 GAGGL--TYGTQNALPLAIDSTSGNATYVFPALP 578
Db 771 SAGHGSGFGTKM-----GDYVAVALP 792

RESULT 13

S00943

Glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - *Acinetobacter calcoace*C/Species: *Acinetobacter calcoaceticus*

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C/Accession: S00943

R/Cleton-Jansen, A.M.; Goosen, N.; Ogde, G.; van de Putte, P.

Nucleic Acids Res. 16, 6228, 1988

A/Title: Nucleotide sequence of the gene coding for quinoprotein glucose dehydrogenase F

A/Reference number: S00943; MUID:88289368; PMID:3399393

A/Accession: S00943

A/Molecule type: DNA

A/Residues: 1-801 <CLE>

A/Cross-references: UNIPROT:P05465; EMBL:X07235; NID:G38711; PIDN:CAA30222.1; PID:G38712

A/Experimental source: strain LMD 79.41

C/Function:

A/Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone

A/Pathway: respiratory chain

C/Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

C/Keywords: oxidoreductase; pyrroloquinoline quinone; respiratory chain; transmembrane F

F/3-35/Domain: transmembrane #status predicted <TM1>

F/3-57/Domain: transmembrane #status predicted <TM2>

F/61-79/Domain: transmembrane #status predicted <TM3>

F/94-108/Domain: transmembrane #status predicted <TM4>

F/118-137/Domain: transmembrane #status predicted <TM5>

F/91,93/Binding site: ubiquinone (Arg, Asp) #status predicted

F/471/Active site: Asp #status predicted

Query Match 12.2%; Score 375.5; DB 1; Length 801;
Best Local Similarity 23.7%; Pred. No. 3.8e-18;

Matches 175; Conservative 89; Mismatches 248; Indels 227; Gaps 38;

QY 6 LRTSAVLLTPAPFAQVTPITDELANP-----AGEINYGROENYRHP 55

Db 121 LSTSLALAIWLVMSIFNDPOEINGEIKTPQPTAQAVPGVARDPAPAGRTQAGVRYSP 180

QY 56 LTTTADNVGLOVVARGWAGAV-----CUTPMIDGVYLANPQDVIOAD 104

Db 181 LKQINDONVADLVAMT--LRTGDLKTDNDSGETTQVPIKIGNMFLCTHAGQLIAD 238

QY 105 AQTGDLIWEHRRLPAVALTNAQGRKGVAYI-----GTSL-----Y 142

Db 239 PATGKEKRRPDKTKDSF--CHLTGRGVYDANNTEFATLSQSKSSSTQCPKRVF 296

QY 143 FSSWDNLILADWETGVNPFVVERSGEDGLTS-----NTGPIVANG--VI 187

Db 297 VPVNDGLVANADTGACTDF--GQNGVNLQEFMPYAYGGINPISPGVLTGSTIV 352

QY 188 VAGS--TCQYS---PYGCFISGHSATGEELM-----RNHFIPQGEGETWGNDE 235

Db 353 IAGSVTDNYSNKEPSG--VIRGYDVTGKLLMVFDTGADBN--AMPGE-----GTTF- 401

QY 236 ARMTGTWGGTTPVNTLVFYSTGVGRASFTQRPQGTLYG-----NTTRA 285

Db 402 VHSNPAWPLADAKLIV--YVPTGV-----GTP--DIWGDRTLEKERYANSMLA 450

QY 286 VRPDTGEIWRHQTLPFDNNDQCTEFEMVANYDVQPSAEMEGIRANP--NATGERR- 342

Db 451 INNSTGLVNFQTHHDLMDVPSQPLAD:KNNAQGTVPALIVLTGKNAFVLDRLN 510

QY 343 -----VLTGAPCKTGTMSS-----FDAASEFLMA----- 367

Db 511 GQPIVPTKRPVQTVKRGQTKGEFYKTPQPSDNLNAPQDLTKDMWGATMLDQLMC 570

QY 368 -----RDNTYNTMTASIDETGLVTVNBDVALKEIDVEYDQPTFLGDRDSSALNDP--- 420

Db 571 RVSEKRLNVDGIYTPSENGTL-----VFPGLGVFEWCGMSVNPDRQV 614
QY 421 -----TGIVF--LPLN--NACVDIVAVDOEFSLADY---YNTSATYAKLAP--GF---E 461
Db 615 AVNMPILGPFYSRLIPADPNRAQPAKAGAGTEGQGVYGVYGVLSAFLSPILGPKCP 674

RESULT 14

B98314

probable quinate dehydrogenase [imported] - *Agrobacterium tumefaciens* (strain C58, CereorC/Species: *Agrobacterium tumefaciens*

C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004

C/Accession: B98314

R/Goodner, B.; Hinkle, G.; Galtung, S.; Miller, N.; Blanchard, M.; Quirollo, B.; Goldman,

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2323-2328, 2001

A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*

A/Reference number: B97359; MUID:21608551; PMID:11743194

A/Accession: B98314

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-809 <KUR>

A/Cross-references: UNIPROT:Q8UML9; GB:AE007870; PIDN:AAK90036.1; PID:G15160013; GSPDB:G

C/Genetics:

A/Map position: linear chromosome

C/Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 12.2%; Score 374.5; DB 2; Length 809;

Best Local Similarity 23.2%; Pred. No. 4.5e-18;

Matches 168; Conservative 80; Mismatches 238; Indels 239; Gaps 31;

QY 19 PAFPAQVTPITDELANPAGEMWINGROENYRHSPLTQITADNVGLOLVW----- 71

Db 158 PAALTPVKEPTEQ-----KMEHYGNTSGGSRFVALDEITRDNIKLEAVMTYHTGDT 210

QY 72 --ARGMAGAVQVTPMTHDGMVYLANPQDVIOALDAGT-----DLWEHRRL 118

Db 211 PISPGANGABDQPTPLQVGVDTVFLCTPHNNVIALDADTGETKMAEINSSKSSVMRCRL 270

QY 119 -----PAVALTNAQGRKGVAYLGTSLYFSSWDNLILADMET--- 157

Db 271 AVPAKAPLQPTAPSGSTPVAAVAAG-----ALQGRILMNTVIAELIALDADTGA 324

QY 158 -----GQVFDVERSGEDGLTSNTTGPVANGVAVAG---STQYSPYCGFISGHD 206

Db 325 CPDFGTNGRVDLKIGMDADPDPQVYVLSAPTLACTTVVAGRIADNVQVDPGAVKRGFD 384

QY 207 SATGEELMRNH-----IPQGEGETWGNDEFAEMTGWGQITVDPVNTLVF- 256

Db 385 VVTGSLRKA--FDGQNDITLCPPGQTYRSTN-----VMSMAYDDDLNIVFM 433

QY 257 -YGSTGVGPASETQRTGGTLYGTNTRFAVRPDTGEIWRHQTLPFDNNDQCTEFEMV 315

Db 434 PVGSPVDLYGATR--TPLDHKYGA--SMLALDATTGKEKVVQTVHNDLMD----- 481

QY 316 ANVDVQPSAEMEGIRANPNATGERVNLGAPC-----XTGTMWSTDAASG-----EF 364

Db 482 FDVMPQPS-----FVDFPRADG---TSVPALVGTGAGQGLVLDRAITGQPLTYEE 529

QY 365 LWARDNTYN--MTASIDETGLVTVN-----DAVLKEL---DVEYD--- 401


```

Db      530 VSVKANIPIEPEYAPTOPRSVGMPOIGAQTITEADMWGATPFDOQLCRIFAKGMRKEGLY 589
QY      402 -----VCPFLGGRDMSAALNPDTGIYFLPLNNACDIAMVDFESA-LDYVNTS 451
Db      590 TAPGTDLSLAFPGSLGGNMWGLSTDPTNTIIFVN-----DRLGLWIEKKEAA 638
QY      452 ATAKLAPGFENM-----GRIDAIDISTGRTL 477
Db      639 PTKVAASGSESVNTGMGVPMKGTPEYAVNKNRFLSALGIPCOAPPYGTMTAIDMKTQGIK 698
QY      478 WSAERPAANY-----PV-----LSTAGGVFENGCT-DRYFRALSGETG 515
Db      699 WGV--PVGIVEDTGPLGIKMLPIPIGMPTLGGTLATOGGLVFLAGTQDYTLAFTDTATG 756
QY      516 ETLMOARLATVATGQAISYE--LDGVQYIAIGAGGLTYGTQLNAPLAIDSTSVGNAY 573
Db      757 KEVWKARLPVSGCGGPMYSKPKTKGQYVVISAGG-----ARQSPDRGDYVI 803
QY      574 VFALP 578
Db      804 AYALP 808

```

RESULT 15

```

A12968
hypothetical protein Atu3354 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C/Accession: A12968
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001.
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A>Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193
A/Accession: A12968
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1809 <KIR>
A/Cross-references: UNIPROT:Q8UAL9; GB:AE008689; PIDN:AAU44167.1; PID:G17741742; GSPDB:G
A/Experimental source: strain C58 (Dupont)
C/Genetics:
A/Gene: Atu3354
A/Map position: linear chromosome
C/Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

```

```

Query Match      12.24; Score 374.5; DB 2; Length 809;
Best Local Similarity 23.24; Pred. No. 4.5e-18;
Matches 168; Conservative 80; Mismatches 238; Indels 239; Gaps 31;

```

```

QY      19 PAFAOYTPITDELALNPAGEMINYGROENYRHSPLQITADNVGQLQLV----- 71
Db      158 PAALTPEKDETBQ-----KMEHYGNISGSRFVALDEITRDNIKLEVAITYTGTDT 210
QY      72 --ARGMEAGAVQYTPMIDGVMYLANGVIALDAQTG-----DLTMEHRQL 118
Db      211 PISFGANGADDOCPFLQVGDVTFLCFPHNVVIALDADTGTEKKWAKINSKSVWRCGL 270
QY      119 -----PAVALINAGDKRGVALYGTSLYFSSMDNHLIALDMET--- 157
Db      271 AYFDAPAKLPKOPTAPGSTPYTAAVVAEG-----ALCQRRLMNTINAEILADADTGA 324
QY      158 -----GQVVPVERSGEGSLTNTTGPVANGVIVAG---STCOYSPYCGFISGD 206
Db      325 CPDEGTNGRVDLKIOMGDADPQYVILSAPTLAGTIVVGGRIADNVQYDMPCGVNRGFD 384
QY      207 SATGEELMRHNF-----IPOGEGEDETWGNDFEARMGTGWGQITYPYTLVF- 256
Db      385 VVTGEIARMA--FDPGNDITKLPFGQTYTRSTPN-----VWASMSYDPDLNTVFM 433

```

```

QY      257 -YGTGTVGASFTQGTGEGTILYGTNTRPVAVRPDTEIWRHQILPRDNMQECTEFEMV 315
Db      434 PVGSPSVLDYGAIR--TPLDHKYGA-SMLADATATGEREKVYQTVHNDLMD----- 481
QY      316 ANVDVQPSAEMGLAIPNATGERRVLTGAPC-----KGTWMSFPAASG-----EF 364
Db      482 FVPEMQPS-----FVDEPKADG-----TSVPALVGTGAGQLVLRATQQLTYVEE 529
QY      365 LVARDTNTN-----MIASIDETGLTVNE-----DAVLEKEL--DVEYD--- 401
Db      530 VSVKANIPIEPEYAPTOPRSVGMPOIGAQTITEADMWGATPFDOQLCRIFAKGMRKEGLY 589
QY      402 -----VCPFLGGRDMSAALNPDTGIYFLPLNNACDIAMVDFESA-LDYVNTS 451
Db      590 TAPGTDLSLAFPGSLGGNMWGLSTDPTNTIIFVN-----DRLGLWIEKKEAA 638
QY      452 ATAKLAPGFENM-----GRIDAIDISTGRTL 477
Db      639 PTKVAASGSESVNTGMGVPMKGTPEYAVNKNRFLSALGIPCOAPPYGTMTAIDMKTQGIK 698
QY      478 WSAERPAANY-----PV-----LSTAGGVFENGCT-DRYFRALSGETG 515
Db      699 WGV--PVGIVEDTGPLGIKMLPIPIGMPTLGGTLATOGGLVFLAGTQDYTLAFTDTATG 756
QY      516 ETLMOARLATVATGQAISYE--LDGVQYIAIGAGGLTYGTQLNAPLAIDSTSVGNAY 573
Db      757 KEVWKARLPVSGCGGPMYSKPKTKGQYVVISAGG-----ARQSPDRGDYVI 803
QY      574 VFALP 578
Db      804 AYALP 808

```

Search completed: November 18, 2004, 05:20:20
Job time : 46 secs

This Page Blank (uspto)